

Issues in Multilingual Frame Semantics: Comparability of frames

Universität Leipzig, October 10 & 11, 2018

Organiser: Oliver Czulo, Institute for Applied Linguistics and Translatology

Presenters

Hans C. Boas (University of Texas in Austin, USA)

Semantic frames as a universal metalanguage?

Oliver Czulo (Universität Leipzig, Germany)

Some thoughts on methods for comparing frames

Benjamin Lyngfelt (Universität Göteborg, Sweden)

Applying English frames to Swedish: Different conditions for multilingual FrameNets and Constructicons

Kyoko Ohara (Keio University, Japan)

Comparability of frames in bilingual children's books

Amanda Patten (University of Birmingham, UK)

Towards a comprehensive Constructicon of English: Bringing together COBUILD Grammar Patterns and FrameNet

Miriam Petruck (International Computer Science Institute, Berkeley, USA)

Issues in aligning frames in the Multilingual FrameNet project

Josef Ruppenhofer (IDS Mannheim, Germany)

Polarity sensitivity in translation

Paul Sambre (KU Leuven, Belgium)

Marriage and 'Ndrangheta: Reframing the role of mafia women in an Italian crime syndicate

Jennifer Sikos (IMS Stuttgart, Germany)

Embeddings and cross-lingual frames

Tiago Torrent (Universidade Federal de Juiz de Fora, Brazil)

Building the GregBot, or how a domain-specific multilingual FrameNet can be turned into a personal travel assistance

Peter Uhrig (Universität Osnabrück, Germany)

The Distributed Little Red Hen Lab - A Community for Research into Multimodal Communication

Alexander Ziem (HHU Düsseldorf, Germany)

The German Constructicon: Status quo and perspectives



UNIVERSITÄT
LEIPZIG

Gefördert durch



Deutsche
Forschungsgemeinschaft

Programme for October 10 & 11

Part 1: Wednesday morning		
10:00-10:30	Oliver Czulo	Introduction Some thoughts on the comparison of frames
10:30-11:00	Hans Boas	Semantic frames as a universal metalanguage?
11:15-11:45	Kyoko Ohara	Comparability of frames in bilingual children's books
11:45-12:15	Miriam Petruck	Issues in aligning frames in the Multilingual FrameNet project
12:15-13:00	<i>Discussion</i>	<i>On the comparability of frames</i>
Part 2: Wednesday afternoon		
14:30-15:00	Paul Sambre	Marriage and 'Ndrangheta: Reframing the role of mafia women in an Italian crime syndicate
15:00-15:30	Sebastian Padó & Jennifer Sikos	Embeddings and cross-lingual frames
15:45-16:15	Tiago Torrent	Building the GregBot, or how a domain-specific multilingual FrameNet can be turned into a personal travel assistant
16:15-17:00	<i>Discussion</i>	<i>Data-driven methods for comparing frames</i>
Part 3: Thursday morning		
09:30-10:00	Alexander Ziem	The German FrameNet and Constructicon Project: Status quo and perspectives
10:00-10:30	Kristian Blenselius & Benjamin Lyngfelt	Applying English frames to Swedish: Different conditions for multilingual FrameNets and Constructicons
10:45-11:15	Amanda Patten & Florent Perek	Towards a comprehensive Constructicon of English: Bringing together COBUILD Grammar Patterns and FrameNet
11:15-11:45	Josef Ruppenhofer	Polarity Sensitivity in translation
11:45-12:30	<i>Discussion</i>	<i>On the relation of constructions and frames</i>
Part 4: Thursday afternoon		
14:00-16:00	<i>Discussion</i>	<i>Putting it all together; German FrameNet?; funding opportunities for collaboration</i>

Presentations

=== Wednesday, 10:00 – 13:00 =====

Some notes on the comparison of frames

Oliver Czulo, Universität Leipzig
czulo@uni-leipzig.de

In (Čulo 2013, Czulo 2017), I devise a first draft of a model of analysis for translation which is based on Construction Grammar and Frame Semantics. The primary hypothesis of this *primacy of frame* model is that in theory, for every source language frame, the corresponding target language frame should appear in the target text. I. e., ideally there would be a one-to-one mapping of frames.

It does not come as a surprise that this is often not the case. The questions around this are: What are the factors which guide such a “deviation” from a one-to-one-mapping? When can we still speak of a “semantic similarity” between original and translation? As for the former question, a variety of factors is assumed: differences in mappings from lexical units to frames, cultural differences or the presence of grammatical constructions in the source language which cannot be easily reproduced in the target language and may lead to further shifts, including semantic shifts in the way the message is constructed. As for the latter question, a small number of examples is discussed in (Čulo 2013), in which frame-to-frame relations can be used to show that semantically, the original and the translated message can be shown to be closely related.

The question which is not tackled in these contributions, however, is that of how “comparable” frames are: Is it possible to e. g. use the frame-to-frame relations from the Berkeley FrameNet (Fillmore et al. 2003) to describe frame relations in German? And in how far can we say of two frames that they are comparable?

In this talk, I will present some works such as (Boas 2001, Ohara 2009) which have focussed on the comparison of frames and will analyse along which dimensions frames were compared. I will also pose the question of the level of granularity on which frames are compared, arguing that a single, abstract frame “structure” can be concretised in various ways within a language community.

References

- Boas, Hans C. 2001. ‘Frame Semantics as a Framework for Describing Polysemy and Syntactic Structures of English and German Motion Verbs in Contrastive Computational Lexicography’. In *Proceedings of Corpus Linguistics 2001*, 64–73.
- Čulo, Oliver. 2013. ‘Constructions-and-Frames Analysis of Translations: The Interplay of Syntax and Semantics in Translations between English and German’. *Constructions and Frames* 5 (2): 143–67.
- Czulo, Oliver. 2017. ‘Aspects of a Primacy of Frame Model of Translation’. In *Empirical Modelling of Translation and Interpreting*, edited by S. Hansen-Schirra, Oliver Czulo, and Sascha Hofmann, 465–90. Translation and Multilingual Natural Language Processing 6. Berlin: Lang-Sci Press.
- Fillmore, Charles J., Christopher R Johnson, and Miriam R L Petruck. 2003. ‘Background to Framenet’. *International Journal of Lexicography* 16 (3): 235–250.
- Ohara, Kyoko Hirose, Seiko Fuji, Toshio Ohori, Ryoko Suzuki, Hiroaki Saito, and Shun Ishizaki. 2009. ‘Frame-Based Contrastive Lexical Semantics in Japanese FrameNet: The Case of risk and kakeru’. In *Multilingual FrameNets in Computational Lexicography*, edited by Hans C. Boas, 163–82. Mouton.

Semantic frames as a universal metalanguage?

Hans Boas, University of Texas, Austin
hcb@mail.utexas.edu

The past twenty years have seen an ever-growing number of studies in Cognitive Linguistics that use large-scale corpora and relevant tools to investigate a range of different linguistic phenomena. Using corpora for linguistic description and analysis has allowed researchers to move away from what Fillmore (1989) famously labeled “armchair linguistics” to a more empirically-based rigorous methodology similar to what Hanks (2014) calls “empirical linguistics” (as opposed to “speculative linguistics”). This talk discusses what happens when a theory that was devised before the advent of what is today known as corpus linguistics, and which is rooted in the ideas of Cognitive Linguistics, namely Frame Semantics (Fillmore 1982), is confronted with naturally occurring data in linguistic corpora. The goal of this talk is to show how semantic frames can be employed as an empirically based meta-language for linguistic analysis that cannot only be used for lexical phenomena, but also for many types of syntactic and phraseological phenomena.

To achieve this goal, this talk focuses on the role of semantic frames as a meta-language for linguistic description and analysis. Of particular interest here is the question of how semantic frames are defined, refined, and finally used to determine what types of corpus data should be extracted and annotated. In the FrameNet workflow, the definitions of semantic frame elements are used by annotators to determine which constituents in a sentence should be categorized as particular frame elements and then annotated as such. This means that semantic frames and their frame elements can be regarded as a corpus-based meta-language. The paper closes with a discussion of how semantic frames are organized in a hierarchy and how they differ from the Natural Semantic Metalanguage (Goddard 2012).

References

- Fillmore, C. J. (1989). Grammatical construction theory and the familiar dichotomies. In *North-Holland Linguistic Series: Linguistic Variations* (Vol. 54, pp. 17-38). Elsevier.
- Goddard, C. (2012). Semantic primes, semantic molecules, semantic templates: Key concepts in the NSM approach to lexical typology. *Linguistics* 50(3), 711-743.
- Hanks, P. (2014). *Lexical Analysis*. Cambridge, MA: MIT Press.

Comparability of frames in bilingual children’s books

Kyoko Hirose Ohara, Keio University
ohara@hc.st.keio.ac.jp

One of the most important research questions in multilingual frame semantics involves establishment of comparability between semantic frames from different languages. This paper aims to define a common theoretical basis for multilingual frame semantic research, by reporting on an ongoing research project, in which Japanese and English bilingual children’s books are annotated with semantic frames (cf. Ohara & Okubo 2018).

Since readers of children’s books are in the process of acquiring their first languages, lexical items and grammatical constructions appearing in children’s books can be regarded as those acquired relatively early by children and thus cognitively basic. It is hypothesized that the semantic frames that are shared by Japanese and English versions of children’s books are cognitively basic and are thus candidates for comparable frames between the two languages.

This paper also investigates influences of factors that may interfere with assignment of same or similar frames in Japanese original texts and their English translations. The *primacy of the frame* hypothesis states that ideally, the frame of the original is matched by the frame of the translation (Czulo 2013: 143). There are, however, two possible factors that may override this principle in the case of frame agreements between Japanese original texts and their English translations of children's books. One is cultural differences that cause establishment of different semantic frames in the two languages; and the other is differences in grammatical constructions preferred by the two languages.

To summarize, the paper aims to contribute to establishment of comparability of frames over languages and cultures by investigating semantic frames in bilingual children's books.

Multilingual FrameNet: Alignment?

Miriam R. L. Petruck, ICSI Berkeley
miriamp@icsi.berkeley.edu

The goal of the Multilingual FrameNet (MLFN) project is to “align the databases for different languages, both at the level of semantic frames and at the level of lexical units”¹. Additionally, that project embarked upon another task, specifically the annotation in several different languages of a text to pursue determining whether and how these languages “frame” the text in the same ways (e.g. Torrent et al 2018), i.e., whether or not the words in the languages under consideration (e.g. English and Brazilian Portuguese) evoke the same frames.

The work of the MLFN project necessitates questioning the notion of alignment in the context of that project. This presentation poses the following question: what does alignment mean for MLFN? Among its goals is to address the conundrum of aligning frames and lexical units among the languages that adopted FrameNet's (Ruppenhofer et al. 2016) frames and lexical units for their own development, notwithstanding Gilardi and Baker's (2018) plan to develop an alignment with the introduction of a new relation of similarity between frames “combining structural and distributional similarity” (2018: 13). Such a frank discussion of this quandary among researchers with first-hand experience in the development of FrameNets for English and other languages is bound to shed light on the issue and ultimately prove useful for the larger effort, and to the extended FrameNet community.

References

- Gilardi, L. and Collin F. Baker. 2018. Learning to Align across Languages: Toward Multilingual FrameNet. In Proceedings of the International FrameNet Workshop 2018: Multilingual FrameNets and Constructicons, T. Torrent, L. Borin and C. F. Baker (eds.), p. 13-22. Co-located with LREC 2018, Miyazaki, Japan.
- Ruppenhofer, Josef, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, Collin F. Baker, and Jan Scheffczyk. 2016. FrameNet II: Extended Theory and Practice, The Book.
- Torrent, T., M. Ellsworth, C. F. Baker, and E. Matos. 2018. The Multilingual FrameNet Shared Annotation Task: A Preliminary Report. In Proceedings of the International FrameNet Workshop 2018: Multilingual FrameNets and Constructicons, T. Torrent, L. Borin & C. F. Baker (eds.), p. 62-67. Co-located with LREC 2018, Miyazaki, Japan.

1 https://www.nsf.gov/awardsearch/showAward?AWD_ID=1629989

=== Wednesday, 14:30 – 17:30 =====

Marriage and ‘Ndrangheta: Reframing the role of mafia women in an Italian crime syndicate

Paul Sambre, MIDI, University of Leuven, Belgium (paul.sambre@kuleuven.be)

This talk examines the FrameNet notion of *marriage/matrimonio* against the background of an Italian (and limited German) corpus of videos about mobster wives (Borromeo 2015, Rai 2015, Sabella 2017) in order to specify how the English FrameNet semantic metalanguage about Frames and Lexical Units such as family and marriage, as an aggregate and kinship relation between spouses may be both used and descriptively extended based on specific cultural situations in other non-English speaking countries such as Italy. Whereas English FrameNet provides positive but rather limited corpus descriptions of the *marriage* lexical unit in terms of establishing family or personal relationships between man and woman, our corpus-based approach allows to describe the specificity of criminal marriages in the Calabrian ‘*Ndrangheta* mafia in terms of strategic alliances between criminal groups well beyond the traditional legal and moral positive values imposed by civil marriage on spouses, now involving participation in secret illegal activities, the reversal of traditional soft feminine roles such as raising children and education to future family life, here in other terms, e.g. as when children are expected to become future mafia members ready to kill. In this world, women may be either possible victims of mafia repression (e.g. in forced marriages, or after so-called perceived adultery) or perpetrators in mafia wars (e.g. as instigators of shootouts, as go-betweens with their imprisoned husbands or initiators of peace processes between rival Mafia clans, eventually leading to their special status as *sister of the code of the silence*). Mafia marriage in other words provides new conceptual content for the definition and traditional categories of equal rights between partners, the obligation of assistance, and the instruction of children and the presumed subordinate role of women in the crime syndicate. Simultaneously, we move beyond the stereotypes of mafia marriage to show how some brave women, such as Maria Concetta Cacciola, Lea Garofalo or Giuseppina Pesce, enter anti-mafia (Sambre 2017) and sometimes die to defy the mafia for love of their husband or children, after becoming collaborators of justice, regardless of the risks such rebel attitude entails (dalla Chiesa 2012, Ursetta 2016). Since the ‘*Ndrangheta* now crosses borders of countries and cultures, we show a short German case of Maria G. (Schraven 2017), who now lives under German police protection, to show how Italian mafia and anti-mafia marriage framing becomes relevant to other languages than Italian (Sambre 2018).

References

- Boas, H. C. 2009. Semantic frames as interlingual representations for multilingual lexical databases. In H.C. Boas (ed.), *Multilingual FrameNets in Computational Lexicography: Methods and Applications*. 59-100. Berlin: Mouton de Gruyter.
- Borromeo, Beatrice. 2015. *Lady ‘Ndrangheta*. Newsweek and Daily Beast.
- dalla Chiesa, Nando. 2012. *Le Ribelli*. Storie di donne che hanno sfidato la mafia per amore. Milano: Melampo.
- Iantosca, Angela. 2013. *Onora la madre*. *Storie di ‘Ndrangheta al femminile*. Soveria Mannelli: Rubbettino.
- Rai. 2015. *A sangue freddo*. La storia siamo noi. Roma. Rai. (2 episodes, about 40 min. each): 9. O cu nnui, o cu iddi. La storia Maria Concetta Cacciola. | 10. La storia di Lea Garofalo.
- Sabella, Marzia. 2017. *Donne di mafia*. Custodi dell’omertà o nuove protagonist della violenza mafiosa. Catania: Auditorium del Monastero dei Benedettini. <https://www.youtube.com/watch?v=AyiqNRQFdEQ> (5 May 2018)

- Sambre, P. 2017. The Multimodal Representation of Italian Antimafia Discourse: Foregrounding Civil Resistance and Interlocution in Two Global English Video Reportages. *I-LanD, Identity, Language, Discourse* 2017(1): 38-59.
- Sambre, P. 2018/fc. Les nouvelles mafias italiennes entre traduction journalistique et analyse du discours critique. Une étude de cas interdiscursive et multimodale sur Mafia Capitale dans le quotidien De Standaard. In: Oster, C., Caliendo, G. (eds.), *Traduire la criminalité. La représentation discursive du crime organisé*. Lille: Septentrion.
- Schraven, David, Meuser, Maik. 2017. *Maria und die Mafia. Eine Kronzeugin packt aus*. DocuCorrectiv/RTL
- Ursetta, Umberto. 2016. *Vittime e ribelli. Donne di 'ndrangheta da Lea Garofalo a Giuseppina Pesce*. Cosenza: Pellegrino.

Embeddings and cross-lingual frames

Jennifer Sikos & Sebastian Padó, IMS Stuttgart
jen.sikos@gmail.com, pado@ims.uni-stuttgart.de

It is a standard assumption in the theory of Frame Semantics that frames will apply across languages. However, even in closely related languages, there exist well-attested cases of frame divergence arising from cross-lingual differences in predicates, semantic roles, and syntactic patterns. In this talk, we explore what insights computational methods can provide in terms of cross-lingual frame applicability.

Our method builds embeddings (vector representations) from frame-annotated corpora for English (FrameNet) and German (SALSA) to represent the ‘meaning’ of frames and frame-evoking predicates based on their usage contexts. This enables us to compare the representations of frames and predicates from English and German to determine how similar their semantic behavior is across the two languages. We compare our frame embedding results against previous observations from the literature. We are able to confirm some previously observed divergences and identify new divergences which are partly motivated by actual differences between the languages, partly by different annotation choices, and partly by corpus differences.

As more FrameNets and frame-annotated corpora become available for different languages, we believe that embedding-based methods can become useful tools, both by expediting the creation of frame-semantic resources for new languages as well as guiding research effort towards frames that might require language-specific alterations.

Building the GregBot, or how a domain-specific multilingual FrameNet can be turned into a personal travel assistant

Tiago Torrent, Universidade Federal de Juiz de Fora
tiago.torrent@ufjf.edu.br

In this talk I address how the Multilingual Knowledge Base for Tourism and Sports (m.knob), developed by FrameNet Brasil, is used in a chatbot providing recommendations of tourist attractions and activities for travelers. The m.knob database currently features more than 5,000 lexical units for three languages - Brazilian Portuguese, English and Spanish - distributed in approximately 300 interlingual frames. Specifically, I demonstrate how the addition of three different kinds of relations to the basic framenet-like structure of the m.knob database allows the system to provide semantic

representations of user's inputs and match them to those produced for tourist attractions based on their descriptions and on comments about them posted to online platforms.

The first of this relation holds between a Frame Element in an eventive frame and the entity frame that defines such FE. This reaction allows for the linkage between a participant in a given conceptual system and another conceptual system elaborating on it. In practical terms, it enables the recommendation system to infer which kinds of entities should be recommended to the user when she informs the system what she wants to do. For example, if the user inputs (1) to the system, the chatbot can recommend restaurants.

(1) I want to eat something.

The second relation represents metonymic substitutions inside entity frames. They are posited between a given core frame element and a non-core element that can be used to refer to the latter. It is used to link indirect references to entities to be recommended by the system. For instance, given the input in (2), the chatbot could recommend pizza and pasta places.

(2) I want to eat Italian.

The last type of relations added to the m.knob database are ternary qualia reactions between LUs. Those are defined in terms of one of the four original qualia relations - constitutive, formal, telic and agentive - that is further specified in terms of a frame in the m.knob database. This approach allows the chatbot to use closely connected LUs, which do not evoke the same frame, as means to interpret indirect references by the user. At the same time, it removes the need to use an ad hoc backbone ontology in m.knob, since the very frames that compose the database are used as specifiers to more general - and, therefore, otherwise less informative - lexical relations.

=== Thursday, 9:30 – 12:30 =====

The German FrameNet and Constructicon Project: Status quo and perspectives

Alexander Ziem, HHU Düsseldorf
ziem@phil.uni-duesseldorf.de

The talk aims at introducing objectives, methods and procedures guiding the German Constructicon Project (GCon) hosted at the University of Duesseldorf (Boas/Ziem in press; Ziem/Boas 2017; Ziem in press).

The presentation is divided into three parts:

(1) *Annotation categories*. The first part introduces the annotation categories applied. Based on the Berkeley FrameNet Constructicon (Fillmore et al. 2012), they include (a) construction evoking elements (CEE); (b) construction elements (CE); and (c) constructs licensed by the construction. In addition, we also annotate so-called “correlated elements” (CorE), that is, strings of words enhancing, or supplementing, a (semantic, pragmatic, discourse-functional, syntactic) property of the target construction.

(2) *Empirical procedure*. Using a great variety of construction families peculiar to German, such as exclamative constructions, reduplication constructions, comparative constructions, negation construction, among others, the second part illustrates the empirical procedure specifically developed for the constructicographic project. The procedure essentially comprises (a) subcorporation and a preliminary analysis; (b) syntactic parsing (using TreeTragger and the Berkeley Parser trained with

German data); (c) semantic annotation with WebAnno; (d) semi-automatic constructional analysis (with the help of a tool called Construction Analyzer); and (e) compilation of construction entries in the GCon database.

(3) *Compiling and storing construction entries.* Finally, based on the empirical results achieved, the third part demonstrates how pieces of constructional information are compiled and stored in a web-based database. By doing so, I will also take a closer look at the general architecture of the German Constructicon. Specifically, I demonstrate to what extent the framework is in line with other ongoing construction projects, most prominently the Swedish Construction (Lyngfelt 2012), the Brazilian Portuguese Constructicon (Torrent et al. 2014) and the Japanese Constructicon (Ohara et al. 2014).

References

- Boas, H. C. & Ziem, A. (in press): Constructing a constructicon for German: Empirical, theoretical, and methodological issues. In: Lyngfelt, B./Borin, L./Ohara, K./Torrent, T. T. (Hrsg): Constructicography: Constructicon development across languages. Amsterdam: Benjamins.
- Fillmore, C. J., Lee-Goldman, R. R., & Rhodes, R. (2012). The FrameNet Constructicon. In H. C. Boas & I. A. Sag (Eds.), *Sign-based Construction Grammar* (pp. 283–299). Stanford: CSLI Publications.
- Lyngfelt, B., Borin, L., Forsberg, M., Prentice, J., Rydstedt, R., Sköldberg, E. & Tingsell, S. (2012). Adding a constructicon to the Swedish resource network of Språkbanken. In *Proceedings of KONVENS 2012* (pp. 452–461). Vienna.
- Ohara, K. H. (2014). Relating Frames and Constructions in Japanese FrameNet. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC '14)* (pp. 2474–2477).
- Torrent, T., Lange, L. M., Sampaio, T. F., da Silva Tavares, T. & da Silva Matos, E. E. (2014). Revisiting border conflicts between FrameNet and Construction Grammar: Annotation policies for the Brazilian Portuguese Constructicon. *Constructions and Frames* 6.1, 33–50.
- Ziem, A. & Boas, H. C. (2017): Towards a Constructicon for German. In: *Proceedings of the AAIL 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding*. Technical Report SS-17-02, Stanford University, 274-277.
- Ziem, A. (in press): *Tag für Tag Arbeit über Arbeit: konstruktionsgrammatische Zugänge zu Reduplikationsstrukturen im Deutschen*. In: Steyer, K. (ed.): Sprachliche Verfestigung. Wortverbindungen, Muster, Phrasem-Konstruktionen (=Studien zur Deutschen Sprache, 79). Tübingen: Narr.

Applying English frames to Swedish: Different conditions for multilingual framenets and constructicons

Kristian Blensenius & Benjamin Lyngfelt, University of Gothenburg
kristian.blensenius@gu.se, benjamin.lyngfelt@svenska.gu.se

Using frames as a base of comparison for contrastive linguistics (cf. e.g. Boas 2009) plays out somewhat differently for lexicon and grammar. The most striking differences are (a) that lexical comparisons mainly concern meaning/function (and possibly distribution), whereas grammatical constructions involve form to a greater extent and in a more direct manner; and (b) that not all grammatical patterns have a frame-like meaning or function.

Regarding the lexical side of things, we will report on the Swedish part of a pilot project for the Multilingual FrameNet (MFN) initiative, consisting of full text annotation of a TED talk translation.

Using English FrameNet frames for the Swedish annotation turned out to work well, for the most part, partly due to English and Swedish being similar languages, partly due to translationese. Instead, annotation problems mostly had to do with expressions not covered by (the current) FrameNet frames, for instance the greeting phrase at the beginning of the text, whose function is somewhat different from the meanings typical of framenet frames. One possible way to handle such expressions would be through grammatical constructions, in an associated constructicon (cf. Fillmore et al. 2012; Lyngfelt, Borin et al. in press).

Including grammatical constructions in an MFN approach raises the issue of how to represent grammatical structure in a cross-linguistically applicable way, as well as how to handle different kinds of meaning relations. Concerning the latter issue, Lyngfelt, Bäckström et al. (in press) discuss adding other kinds of construction-to-frame relations than the standard evoking relation, whereas Ohara (in press) proposes so-called interactional frames in addition to the semantic frames typical of FrameNet. In our presentation, we will discuss these issues in relation to the Swedish MFN annotation and the Swedish constructicon.

References

- Boas, Hans C. (ed.) (2009). *Multilingual framenets in computational lexicography: Methods and applications*. Berlin & New York: Mouton de Gruyter.
- Fillmore, Charles J., Russell Lee-Goldman & Russell Rhomieux (2012). The FrameNet Constructicon. In H. C. Boas & I. A. Sag (Eds.), *Sign-Based Construction Grammar* (pp. 309–372). Stanford: CSLI Publications.
- Lyngfelt, Benjamin, Linnéa Bäckström, Lars Borin, Anna Ehrlemark & Rudolf Rydstedt (in press). In Lyngfelt, Borin et al. (eds.).
- Lyngfelt, Benjamin, Lars Borin, Kyoko Hirose Ohara & Tiago Timponi Torrent (eds.) (in press). *Constructicography: Constructicon development across languages*. Amsterdam: John Benjamins.
- Ohara, Kyoko (in press). Relations between frames and constructions. A proposal from the Japanese FrameNet constructicon. In Lyngfelt, Borin et al. (eds.).

Towards a comprehensive constructicon of English: Bringing together COBUILD Grammar Patterns and FrameNet

Amanda Patten & Florent Perek, University of Birmingham
a.l.patten@bham.ac.uk, f.b.perek@bham.ac.uk

This paper outlines our proposal to develop a more comprehensive constructicon of English, using the existing resources of COBUILD Grammar Patterns and FrameNet.

The two-volume COBUILD grammar patterns series (Francis et al. 1996, 1998) catalogues the syntactic environment of lexical items in the Bank of English corpus. Over 200 language patterns are listed, along with the set of lexical items (nouns, verbs and adjectives) attested within them. The COBUILD “patterns” (Hunston & Francis 2000) are similar to constructions, in that they are conceptualised as single coherent grammatical units consisting of fixed parts and open slots. However, while each pattern’s entry includes lexical sense information, and the lexical items are sorted into intuitive “meaning groups”, the patterns themselves are not explicitly paired with meaning or with semantic role descriptors. The FrameNet database includes valency information describing how frame elements are encoded in BNC corpus examples (Ruppenhofer et al. 2016). These are matched with the COBUILD patterns, and so provide the semantic component for our constructions.

This paper provides an overview of our methodology, exemplified through a pilot study. We focus in particular on our decision-making process as to the design of the constructional entry (e.g. issues of grammatical description, capturing the division of labour between lexis and construction) and the architecture of the constructional network (e.g. positing relations between constructions). We compare our approach with that taken in construction projects of other languages (see Lyngfelt et al. in press) as well as in the Berkeley FrameNet Construction of English (Fillmore et al. 2012).

References

- Fillmore, C. J., Lee-Goldman, R. R., & Rhomieux, R. (2012). The FrameNet Construction. In Sag, I. A., & Boas, H. C. (eds.), *Sign-Based Construction Grammar*. Stanford: CSLI, pp. 283–322.
- Francis, G., Hunston, S. and Manning, E. (1996). *Collins COBUILD Grammar Patterns 1: Verbs*. London: HarperCollins.
- Francis, G., Hunston, S. and Manning, E. (1998). *Collins COBUILD Grammar Patterns 2: Nouns and Adjectives*. London: HarperCollins.
- Hunston, S. & Francis, G. (2000). *Pattern Grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: Benjamins.
- Lyngfelt B., Borin L., Ohara K. and Torrent T.T. (eds.) (in press) *Constructicography: Constructicon development across languages*. Amsterdam: John Benjamins.
- Ruppenhofer, J, Ellsworth M., Petruck M.R.L., Johnson C.R., Baker C.F. Scheffczyk J. (2016) *FrameNet II: Extended Theory and Practice (Revised November 1, 2016)*. Available at: <https://framenet2.icsi.berkeley.edu/docs/r1.7/book.pdf>

Polarity Sensitivity in translation

Josef Ruppenhofer, IDS Mannheim
ruppenhofer@ids-mannheim.de

As traditionally understood, polarity-sensitive items (PSIs) are expressions with a distribution limited to either affirmative or negative contexts (Giannakidou [2011]). Recent work such as Israel [2011] has emphasized that large numbers of PSIs such as *lift a finger* and *all the time in the world* subserve discourse routines like understatement and emphasis and that lexical-semantic classes need to be invoked when analyzing the properties of PSIs. Ruppenhofer and Michaelis [2016] sharpened this approach by demonstrating the utility of frame and construction-based analysis in extending and sub-analyzing Israel's collection of PSIs.

In this contribution, we take the above findings as a point of departure and focus on the cross-lingual alignment of PSI types and tokens between German and English. On the one hand, we expect to find a significant amount of parallelism because the functional motivation(s) for the grammaticalization of PSIs can be presumed to operate in both languages and because the languages are closely related. On the other hand, it is well known that there are also many idiosyncratic-seeming PSIs in the two languages that have no clear counterpart in the other language (e.g. German *das Gras wachsen hören*). With regard to such cases, we are interested in studying a) if idiosyncrasy is more commonly found with certain PSI subtypes than others and b) how such PSIs are rendered in translation. We will present empirical results based on data from parallel corpora and translation dictionaries.

References

- Giannakidou, Anastasia. Negative and positive polarity items: Variation, licensing, and compositionality. In Claudia Maienborn, Klaus von Stechow, and Paul Portner, editors, *Semantics: An international handbook of natural language meaning*, volume 33. Walter de Gruyter, 2011.

Israel, Michael. The grammar of polarity: Pragmatics, sensitivity, and the logic of scales, volume 127. Cambridge University Press, 2011.

Ruppenhofer, Josef and Laura A. Michaelis. Frames, polarity and causation. Corpora, 11(2):259-290, 2016. doi: 10.3366/cor.2016.0094. URL <https://doi.org/10.3366/cor.2016.0094.1>

=== Thursday, 14:00 – 16:00 =====

INCUBATORS – DISCUSSION: Connecting the Community

Common platforms for collecting and disseminating publications

Oliver Czulo, Universität Leipzig

The Distributed Little Red Hen Lab - A Community for Research into Multimodal Communication

Peter Uhrig, Universität Osnabrück

The Distributed Little Red Hen Lab is a global co-operative of researchers working mainly on multimodal communication. It is not owned by any individual institution and its sole purpose is to connect scholars with different resources, skills, etc. to address the challenging research questions of multimodal communication. While it is centred around the NewsScape dataset, it is certainly not limited to it. Red Hen is not a service provider - it can be compared to a gift economy or a sharing culture where the results of one's own work is shared with other Red Hens. Red Hen is devoted to Open Source (4th time in Google Summer of Code 2018). The presentation will give a brief overview of the data, infrastructure, and tools that are already in place but will also discuss areas that are - as of now - underdeveloped.