# Implementing semantic frames and constructions in GF

Normunds Grūzītis

REMU Retreat 2015

http://remu.grammaticalframework.org/retreat/2015/

# FrameNet

- A lexico-semantic resource based on the theory of frame semantics (Fillmore et al., 2003)
  - A semantic **frame** represents a prototypical, language-independent situation characterized by **frame elements** (**FE**) – **semantic valence**
  - Frames are evoked in sentences by language-specific **lexical units** (**LU**)
  - FEs are mapped based on the **syntactic valence** of the LU
    - The syntactic valence patterns are derived from **FrameNet-annotated corpora** (for an increasing number of languages)
  - FEs are divided into **core** and **non-core** ones
    - Core FEs uniquely characterize the frame and syntactically correspond to verb **arguments**
    - Non-core FEs (**adjuncts**) are not specific to the frame

# Example

| Desiring | |
|---|---|
| Definition: | An Experiencer desires that an Event occur. In some cases, the Experiencer is an active participant in the Event, and in such cases the Event itself is often not mentioned, but rather some Focal_participant which is subordinately involved. |
| Core: | Event, Experiencer, Focal_participant, Location_of_Event |
| Non-core: | Cause, Degree, Duration, Manner, Place, Purpose_of_Event, Reason, Role_of_focal_participant, Time, Time_of_Event |

**BFN** frames and FEs

*want*.v..6412

| Examples | Valence patterns | |
|---|---|---|
| 40 | Event | Experiencer |
| (22) | VPto.Dep | NP.Ext |
| 14 | Experiencer | Focal_participant |
| (10) | NP.Ext | NP.Obj |
| (1) | PP[by].Dep | NP.Ext |

Some valence patterns found in **BFN**

*känna_för*.vb..1

| Examples | Valence patterns | |
|---|---|---|
| 1 | Event | Experiencer |
| (1) | VB.INF.VG | NN.SS |
| 2 | Experiencer | Focal_participant |
| (2) | NN.SS | NN.OO |

Some valence patterns found in **SweFN**

# FrameNet-based grammar in GF

- Existing FNs are not entirely formal and computational
  - We provide a **computational** FrameNet-based grammar and lexicon
- GF, Grammatical Framework (Ranta, 2004)
  - Separates between an **abstract syntax** and **concrete syntaxes**
  - Provides a general-purpose **resource grammar library** (RGL) for nearly 30 languages that implement the same abstract syntax
    - Large mono- and multilingual **lexicons** (for an increasing number of languages)
- The language-independent layer of FrameNet (frames and FEs) – the abstract syntax
  - The language-specific layers (surface realization of frames and LUs) – concrete syntaxes
- RGL is used for unifying the syntactic types used in different FNs
  - FrameNet allows for abstracting over RGL constructors

# Initial aim

- Provide a shared FrameNet API to GF RGL, so that application grammar developers could primarily use semantic constructors
  - In combination with some simple syntactic constructors
  - But instead of comparatively complex constructors for building verb phrases

```
mkCl person (mkVP (mkVP live_V) (mkAdv in_Prep place))
    -- mkCl : NP -> VP -> Cl
    -- mkVP : V -> VP
    -- mkVP : VP -> Adv -> VP
    -- mkAdv : Prep -> NP -> Adv
```

```
Residence                       -- Residence : NP -> Adv -> V -> Cl
    person                      -- NP (Resident)
    (mkAdv in_Prep place)       -- Adv (Location)
    live_V_Residence            -- V (LU)
```

# A FrameNet-based API to GF Resource Grammar Library

Frames | Verbs

vers. 0.9.3

REMU

Damaging
Daring
Death
Deciding
Delimitation_of_diversity
Delivery
Deny_permission
Departing
Deserving
Desiring
Destroying
Detaching
Detaining
Differentiation
Dispersal
Dodging
Dominate_competitor

| | | |
|---|---|---|
| 1. | Sweden | 138 (26.24%) |
| 2. | (not set) | 77 (14.64%) |
| 3. | United States | 73 (13.88%) |
| 4. | Russia | 66 (12.55%) |
| 5. | South Korea | 49 (9.32%) |
| 6. | China | 13 (2.47%) |
| 7. | Latvia | 10 (1.90%) |
| 8. | Japan | 8 (1.52%) |
| 9. | Spain | 7 (1.33%) |
| 10. | United Kingdom | 7 (1.33%) |
| 11. | Ireland | 7 (1.33%) |
| 12. | Canada | 6 (1.14%) |
| 13. | Germany | 5 (0.95%) |
| 14. | France | 5 (0.95%) |

## Desiring

**Desiring_V** : Experiencer_NP → Focal_participant_Adv → V → Clause

- **Eng**: [he]$_{\text{Experiencer}}$ [WANTED] [more]$_{\text{Focal\_participant}}$
  - **aspire**_V_Desiring : V
  - **hanker**_V_Desiring : V
  - **hunger**_V_Desiring : V
  - **long**_V_Desiring : V
  - **lust**_V_Desiring : V
  - **pine**_V_Desiring : V
  - **thirst**_V_Desiring : V
  - **want**_V_Desiring : V
  - **yearn**_V_Desiring : V
- **Swe**: [Roberte]$_{\text{Experiencer}}$ [LÄNGTADE] [hem till Tyskland]$_{\text{Focal\_participant}}$
  - **längta**_V_Desiring : V

**Desiring_V2** : Experiencer_NP → Focal_participant_NP → V2 → Clause

- **Eng**: [you]$_{\text{Experiencer}}$ [WANT] [one]$_{\text{Focal\_participant}}$
  - **covet**_V2_Desiring : V2
  - **crave**_V2_Desiring : V2
  - **desire**_V2_Desiring : V2
  - **fancy**_V2_Desiring : V2
  - **want**_V2_Desiring : V2
  - **yearn**_V2_Desiring : V2

http://grammaticalframework.org/framenet/

# Future work

- Add more languages
  - Cooperation needed
- Separate LU-governed **prepositional objects** from adverbial modifiers (*Adv* vs. *NP* arguments)
- Differentiate syntactic **roles of *VP* FEs** (object vs. adverbial modifier)
- Include shared **non-core FEs** (via a modified comparison algorithm)
- **Align LUs** among languages (e.g. via GF translation dictionaries) √
- Towards FrameNet parsing in GF
  - First, frame labelling
    - FrameNet grammar as an **embedded CNL** in RGL
    - Restrict **LUs to frames** (by using GF dependent types)
  - Later, semantic role labelling (SRL)

# Constructicon

- Somewhere between the syntax and lexicon
- **Lexical units**: *word*-meaning pairs (FrameNet)
  - Incl. fixed multi-word expressions
- **Constructions**: *form*-meaning pairs
  - Each construction contains at least one variable element
  - At least one fixed element? OR Everything "above" the lexicon?
- An example: *make one's <u>way</u>* (WAY_MEANS) [1]
  - Structure: {Motion verb [Verb] [PossNP]}
  - Evokes: MOTION
    - [Theme*They*] {*hacked their way*} [Source*out*] [Goal*into the open*].
    - [Theme*We*] {*sang our way*} [Path*across Europe*].
    - *Hopefully* [Theme*he*]*'ll* {*make his way*} [Goal*to our location*].

# Multilinguality

- Berkeley Constructicon (BCxn)
  - A pilot project (~70 constructions)
- **Swedish** Constructicon (SweCxn)
  - An ongoing project (~300 constructions so far), inspired by BCxn
- Brazilian Portuguese Constructicon, few other constructicons
  - Ongoing projects, inspired by BCxn

- Translation is not always compositional
  - A multilingual constructicon would help to make it compositional ("again") [2]

- *Constructions with a referential meaning may be linked via FrameNet frames, while those with a more abstract grammatical function may be related in terms of their grammatical properties.* [3]

## GF Translator:

Swedish | Clear | ⇆ | English | Colors | Translate

Jag behöver mat till festen.

I need food to the party.

Enter text to translate above

Try Google Translate

1 2 3 4 5 6 7 8 9 10

42.898373

PhrUtt NoPConj (UttS (UseCl (TTAnt TPres ASimul) PPos (PredVP (UsePron i_Pron) (AdvVP (ComplSlash (SlashV2a need_V2) (MassNP (UseN food_N)))
(PrepNP to_Prep (DetCN (DetQuant DefArt NumSg) (UseN party_1_N))))))) NoVoc

## Google Translate:

Swedish | English | Latvian | Detect language | ⇆ | English | Latvian | Spanish | Translate

Jag behöver mat till festen. | I need food to the party.

Jag behöver mat till festen | I need food for the party

## SweCxn:

behöva_något_till_något - *behöver mat till festen*

| | |
|---|---|
| **category** | VP |
| **FrameNet** | Needing |
| **structure** | [behöva1 $NP_1$ till1 $NP_2$ \| VP ] |

[4]

# Why GF?

- Constructions is a mixture of lexical units and syntactic rules – there is no formal distinction between lexical and syntactic functions in GF; it fits the nature of constructicons
- The support for multilinguality
- Constructicon as an embedded grammar
- An extension to the GF FrameNet grammar and lexicon

# Implementation in GF

- **Automatic** normalization and consistency checking
  - Feedback (errors and warnings)
- **Automatic** generation of the abstract syntax
  - For each construction, 1..N functions
    - Alternative/optional variables vs. alternative/optional lexical units
- **Automatic** generation of the concrete syntax (partial)
  - By systematically applying the high-level RGL constructors
    - And limited low-level means (ToDo)
  - Pseudo syntax → Actual syntax
  - Feedback (success and failures)
- **Manual** verification and completion (ToDo)
  - Requires a good knowledge and linguistic intuition of the language (Swe) and, preferably, a corpus; low-level knowledge of RGL

# Abstract syntax

- *behöva_något_till_något*
  - Type: VP
  - Structure: [*behöva*..1 **NP**_1 *till*..1 **NP**_2|**VP**]
  - Detailed description (partial):
    - {cat=V, role=State, lu=behöva..1}
    - {cat=NP, role=Requirement}
    - {cat=P, lu=till..1}
    - {cat=NP|Pn|VP, role=Recipient}

- `fun behöva_något_till_något_VP_1 : NP -> NP -> VP`
- `fun behöva_något_till_något_VP_2 : NP -> VP -> VP`

# Concrete syntax

- Many constructions can be implemented by systematically applying the high-level RGL constructors
  - A parsing problem: which constructors in which order?

| Construction | Elements | Patterns |
|---|---|---|
| behöva_något_till_något_VP_1 | behöva_V NP_1 till_Prep NP_2 | V NP Prep NP |
| behöva_något_till_något_VP_2 | behöva_V NP_1 till_Prep VP | V NP Prep VP |

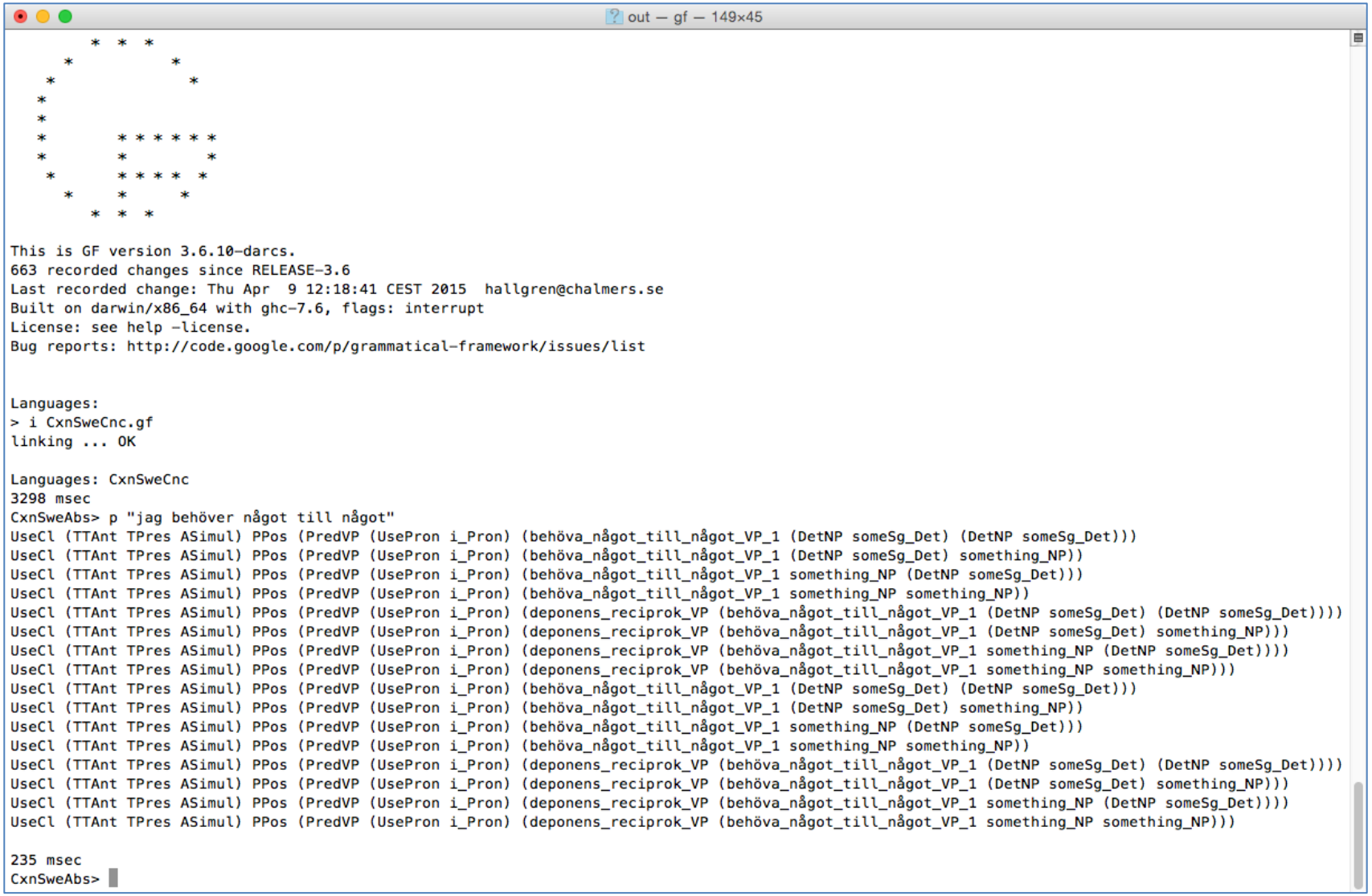
| Pseudocode |
|---|
| `mkVP (mkVP (mkV2 mkV) NP) (mkAdv mkPrep NP)` |
| The parser failed at token VP (no example justifies the *V NP Prep VP* case) |

A simple GF grammar
(might need some manual probs)

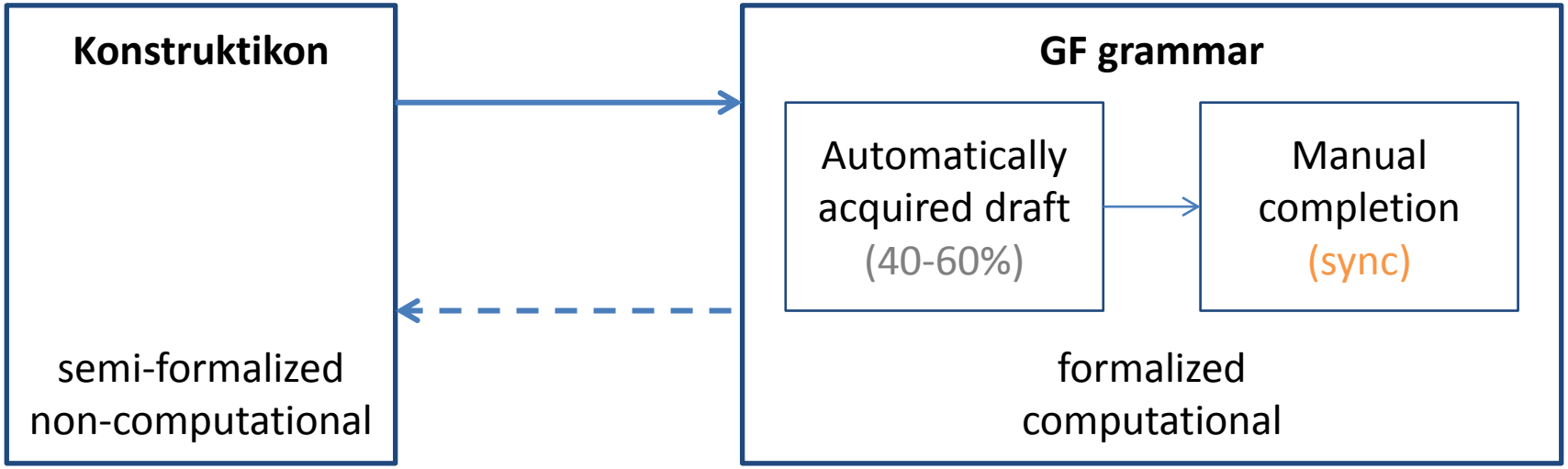
**Final code (by automatic post-processing)**

```
lin behöva_något_till_något_VP_1 np_1 np_2 = mkVP
  (mkVP (mkV2 (mkV "behöver")) np_1)
  (SyntaxSwe.mkAdv (mkPrep "till") np_2) ;
```

# Running example



# Implementation in GF




**ToDo**:

1. Extend and apply the automated approach to all types of constructions in SweCxn
2. Conduct a manual / corpus-based evaluation
3. Write a paper to GEAF and/or some other venue (LREC 2016, ICCG, ...)

- Integration with the FrameNet grammar
- Mapping to the BCxn (a shared abstract syntax)

# References


- [1] Fillmore Ch. J., Lee-Goldman R. R., Rhodes R. The FrameNet Constructicon. In: Boas H. C. and Sag I. A. (Eds.), Sign-based Construction Grammar, Stanford: CSLI, 2012
- [2] From a CLT meeting on constructions [2014/06/10]
- [3] Bäckström L., Lyngfelt B., Sköldberg E. Towards interlingual constructicography. Constructions and Frames, 6(1):9–32. John Benjamins Publishing Company, 2014
- [4] http://spraakbanken.gu.se/eng/resource/konstruktikon/development-version