Construction Annotation Workshop

Michael Ellsworth
ICSI/UCB
June 26, 2017

Outline and Overview

- Introduction to Frames and Constructions

- Annotating Frames and Constructions

- Hands-on session

- Questions and discussion

- Throughout:
  - Finding examples
  - Annotating
  - Evaluating

Outline

- Introduction to Frames and Constructions
  - FrameNet
  - FrameNet Constructicon
  - Joint exercise
- Annotating Frames and Constructions
- Hands-on session
- Questions and discussion

FrameNet

- On-going project
- Goal: demonstrate word meaning and use
- Results:
  - 1.2k semantic **Frames**
  - 1.8k semantic relations between Frames
  - 13k word senses (**Lexical Units**) assigned to frames
  - 200k annotated word-sense instances
    - Semantic roles (**Frame Elements**) and their syntax
    - Other syntax/semantics: relative clause, metaphor, ...

Example Annotation: Forgiveness

- I hope [you$_{Judge/Ext/NP}$] 'll FORGIVE$_{Target}$ [me$_{Evaluee/Obj/NP}$] [for having knocked you down that day$_{Offense/Dep/PPing(for)}$]

- [She$_{Judge/Ext/NP}$] 's FORGIVEN$_{Target}$ [me$_{Evaluee/Obj/NP}$] [for being such an ass with my ultimatum the other day$_{Offense/Dep/PPing(for)}$]

- I think [they$_{Evaluee/Ext/NP}$] will be FORGIVEN$_{Target}$ [for this slight oversight$_{Offense/Dep/PPing(for)}$] .[CNI$_{Judge}$]

  - **Basic valence**: Judge/Ext/NP + Evaluee/Obj/NP + Offense/Dep/PPing(for)

Constructions in FrameNet

- Valences and valence alternations
- Multi-word expressions (LUs with pos *.idio*)
  - *Dog (someone's) steps.idio* in Cotheme frame
- Constructional Null Instantiation
- Relative clauses
  - [The couple_{Partners/Ext/NP/**Ant**}] , [who_{Partners/Ext/NP/**Rel**}] DIVORCED four months ago after 16 years of marriage ...
- (So-called) Raising, small clauses, etc.
  - I 've always THOUGHT [that_{Content/NP/Obj}] [much more undignified than going bald_{Content/Dep/AP}] .

1ˢᵗ Constructicon (Fillmore and the Russells)

- How to represent non-lexical constructions?
  - Many ideas, but still an open question! (and depends on your purpose)
- Can we leverage FrameNet methods?
  - Yes, definitely!
  - But: Lexical Unit must be exploded into **Construct** with constituent **Construct Elements** and optional(!) **Construction Evoking Element**
  - [she$_{Theme}$] {[felt$_{Transitive\_means\_verb}$] [<her$_{Theme}$ way>$_{Way\_expression}$]**Way_means**} [forward$_{Direction}$]
- Product
  - ~1k annotations
  - A few dozen constructions organized into families
  - Method demonstrated for the range from non-lexical (Shared completion) to lexical, but weird (Way-cxns)

Construction Annotation

- Much like FrameNet, but annotates the pieces of a **Construct** and context, as well as valences

- [I$_{individual\_1/Ext/NP}$] 'm$_{Cop}$ {[friends$_{Head\_noun}$]**Be_recip**} [with the Finnish President$_{Individual\_2/Dep/PP(with)}$] .

- [ya$_{Individual\_1/Ext/NP}$] kin be$_{Cop}$ {[partners$_{Head\_noun}$]**Be_recip**} [with me$_{Individual\_2/Dep/PP(with)}$] .

- {[the left leg$_{Before}$] [is$_{Gapped\_portion}$] [forward$_{After}$] [,$_{Punctuation}$] [the right$_{Before}$] [back$_{After}$]**Gapping**}

- http://www1.icsi.berkeley.edu/~hsato/cxn00/21colorTag/

Joint Exercise: Question Augments

- Ex:
  - <u>What the heck</u> are you doing?
  - <u>WTF</u>
  - <u>What in the name of all that is good and holy</u> is a MUD?

  - Further examples?
  - (If we were doing this for real, we'd also do a lit check, of course)

Our task

- Find sentences that show the construction
  - Need diversity!
  - Need accuracy!
- Annotate sentences for:
  - Whatever seems to be important or special for understanding or recognizing them
- Analyze whether we've done a good job
  - Do we cover diversity and exemplify well enough?

Finding the construction

- Tools and corpora:
  - Google
  - CQP or other query tool (cf. COCA corpus)
  - Tree grep, only on parsed corpora
  - Other tools?
- Google is mostly appropriate for constructions with fixed lexical material (which this is)
- Can get more believable results by querying BNC (or other balanced corpus) with tools.

Finding examples: Goals

- Diversity
  - Construct-internal diversity
    - e.g.: Question words, Augments
  - Construct-external diversity
    - Kinds of clauses it occurs in
    - Are there negation requirements?
    - Any other conditions or norms?
- Accuracy
  - What percentage are positive for each search?

Finding examples: Best practices

- Try the easiest appropriate method first
- Use a variety of methods, see what you get
- Pass even low-accuracy search results to annotation stage
- Iterate:
    - You don't need the perfect searches if you can check inadequacies with annotation and review stages!

Annotation

- Which examples to annotate?
  - Start with simple, add more depending on purpose
- What to annotate?
  - Always
    - The pieces that make the construct itself, especially a Construction Evoking Element, if there is one
    - Contextual elements required for construction
  - Usually
    - Construct-external Construction Elements (valences)
    - Other interacting constructions (as for FrameNet annotation)

Annotation Analysis/Review

- Write-up of construction
  - Characterize the way a construction is built
  - Exemplify diversity
  - Give examples that show meaning and use

- Does the sum total of annotation support the analysis and satisfy the needs of your project?
  - If not, add to annotation or even go back to searches
  - Iterate until fully baked and golden brown

Break Time!

Outline and Overview

- Introduction to Frames and Constructions

- Annotating Frames and Constructions

    - Principles

    - Guidelines

    - Group exercises

    - Review of exercises

- Hands-on session

- Questions and discussion

Deeper into Construction Anno

- Principles
  - A complete cline from highly lexical to completely schematic constructions
  - Language is produced by humans! It will always be more complex than we can fully handle, but we try.
  - Do not ignore data!
  - Whatever we need to describe construction!

Guidelines/Recommendations

- Use the right tools and corpora

- Preserve diversity in searches, even if accuracy is low

- Cleanly separate construct-internal and construct-external elements

- Always use your constructionist friends to critique your analysis

- Don't sweat making everything perfect immediately. Iterate!

Exercise: Wh-Exclamatives

- <u>What an idiot</u>!

- <u>How nice</u>!

- I didn't know <u>how good</u> I had it!

- Clearly a construction or group of constructions —these don't match question meaning (they don't ask anything), and their form is wrong for questions. My label gives away some of their presumed function...