Corpus methods for research in pragmatics

Judith Degen, Stanford University

17/08/2016

ESSLLI 2016, Bolzano

“[…] if we stick entirely to artificial examples there is a danger that we will end up with artificial theories.” — Beaver 2010, p. 26

Today

- TGrep2 tutorial
- hands-on project intro: the projection behavior of factive verbs Beaver 2010; Tonhauser 2016; Simons et al to appear

Projection behavior of complements of cognitive factive attitude verbs

The standard story about presupposition

(1) Main clause

She discovered that he lied.

(2) Negation

She didn't discover that he lied.

(3) Antecedent of conditional

If she discovered that he lied, she didn't show it.

(4) Interrogatives

Did she discover that he lied?

The lexical entry of "discover" requires that its complement be presupposed. Karttunen 1974; Heim 1983, 1992; van der Sandt 1992; Schlenker 2008

Problem: fragility of projection

Beaver 2010

*If I **discover** that your work is plagiarized, I will be forced to notify the Dean.*

*Did you learn something about yourself? Did you **discover** that you were stronger than you thought?*

*I haven't tried this with wombats, though, & if anyone **discovers** that the method is also wombat-proof, I'd really like to know!*

Person and projection

Beaver 2010

Person does not fully determine projection. Cases of second person *with* projection:

*Did you ever **realize** that your "mouse pad" makes a great beer coaster?*

*Did you **discover** that this site is filled with "hidden" pages?*

Unexpected projection

Beaver 2010

*You would think I'd know, given I'm an optometrist. But because of the stroke, I am not **aware** that I don't see everything. My wife Robyn noticed that I left food on the left side of the dinner plate. I was still hungry but I thought I had finished the meal.*

Intonation and focus

Beaver 2010; Simons et al 2015

(1) If I discover that your work is [**plagiarized**]$_F$, I will be [**forced to notify the Dea**n]$_F$.

(2) If I [**discover**]$_F$ that your work is plagiarized, I will be [**forced to notify the Dean**]$_F$.

(3) If the T.A. discovers that your work is [**plagiarized**]$_F$, I will be [**forced to notify the Dean**]$_F$.

(4) If the T.A. [**discovers**]$_F$ that your work is plagiarized, I will be [**forced to notify the Dean**]$_F$.

Simons et al's CQ account

Simons et al to appear

Projection of the content of the complement of an attitude verb occurs if the Congruent Question (CQ) of the utterance entails this content.

The CQ for an utterance is a privileged subset of the focal alternative set of the uttered sentence […] which meets the following conditions:
(i) The proposition expressed is a member of the CQ.
(ii) The CQ has at least one additional member.

Example

[JOHN]$_F$ didn't eat the cake.

CQ: domain restricted version of either (1) or (2)
(1) Who ate the cake?
(2) Who didn't eat the cake?

- Claim: (1) is more plausible CQ (for reasons having to do with intonational details, see also Beaver & Clark)
- All alternatives in CQ for (1) entail *that someone ate the cake*

CQ(1) = {John ate the cake, Mary ate the cake, Frank ate the cake, Jane ate the cake}

So why corpora?

“So, while it is true that there are uses of sentences with factive verbs which imply the truth of the propositional complement, there are also uses of sentences with factive verbs which do not. We have little evidence as to whether one case is more frequent or common than the other.”

— Simons et al 2015

Questions

(1) How much **variation** is there in the projection behavior of the complements of cognitive factive verbs?

(2) Which **factors** determine this projection behavior?

(3) Do factors all contribute **independent explanatory power** or are some subsumed under others?

Factors that affect projection

- person *(1st, 2nd, 3rd)*
- intonation / focus *(prominence of verb compared to complement)*
- speaker knowledge *(of complement)*
- speaker intentions
- prior probability of complement
- prior probability of knowing truth of complement (is this the same as speaker knowledge?)
- the verb itself

Factives from Beaver 2010

Cognitive factives

know
realize
discover
notice
recognize
find out
remember
forget
be aware that
be unaware that
admit
intuit

Sensory factives

sense
see
smell
hear
detect
observe

How to proceed?

- choose corpus
  - Switchboard: spoken, parsed, intonation info
- generate and test patterns
- build database

But

“I doubt that there is any general principle that would enable one to predict from the written form of an arbitrary sentence involving a cognitive factive whether the factive complement is presupposed by the author” — Beaver 2010, p. 26

Crowd-sourced judgments

We can't extract everything from the corpus:

- prior probability of complement truth
- speaker intentions
- **projection judgments**

Crowd-sourcing to the rescue!

Next steps

- Today:
  - develop tgrep2 patterns to extract cognitive factive in different environments (negation, antecedent of conditional, questions)
  - develop tgrep2 patterns to extract syntactic features of interest (eg person)
- Tomorrow:
  - think about how to get crowd-sourced projection judgments
  - think about which additional factors we can automatically extract from corpus (eg intonation)

Brainstorming crowd-sourced judgments

- control condition? (ie don't just extract factives, but also clearly non-factives like "believe")

- randomly sample some number of matches and annotate for projection?

- what sort of projection measure? 3-point? does, doesn't, not sure? or slider? or 5-point? should it be does/doesn't or framed in terms of speaker commitment?