Lexical semantics phenomenon: Monotonicity

Upward Monotone (default):
I know I’ll buy a Mac  $\Rightarrow$ I know I’ll buy a computer
subset of

Downward Monotone:
I doubt I’ll buy a Mac  $\Rightarrow$ I doubt I’ll buy a computer
subset of

Downward entailing operators invert the default monotonicity, allowing one to “reason from sets to subsets” [van den Wouden, 1997]

Examples:
He came without cash or cards  $\Rightarrow$ He came without cash
She is too lazy to run  $\Rightarrow$ She is too lazy to run a 10k
I am reluctant to steal money  $\Rightarrow$ I am reluctant to steal SEK

Task: Automatically discover DE operators

Why?

✓ Linguistic importance
DE operators play "an extremely important role in natural language". [van der Wouden, 1997; van Benthem, 1986; Hoeksema, 1988; Dowty, 1994; Sanchez Valencia, 1991]

✓ Textual entailment
TE systems rely on small hand-annotated lists of DE operators. [Nairn et al., 2006; MacCartney and Manning, 2008; Bar-Haim et al., 2008]

✓ Natural language generation
DM inferences induce greater cognitive load than UM inferences [Geurts and van der Slik, 2005]

✓ Prevalence
At least 6% of newswire sentences contain a non-trivial DE operator [Danescu-Niculescu-Mizil et al., 2009]

But how?

梏 No monotonicity-annotated corpora
梏 Not deducible from any public lexical database [Nairn et al., 2006]

Results for Romanian

First time DE ops. are learned for a language other than English!

System output and annotation available at: www.cs.cornell.edu/~cristian/acl2010

Linguistic insight

Ladusaw’s (1980) Hypothesis:
Negative Polarity Items (NPIs) only appear within the scope of downward-entailing operators.

He does not listen anymore vs. *He does listen anymore I doubt you have a clue vs. *You have a clue

Where’s the green? It is wise to try compensating for any excess.
S(x) := fraction of NPI contexts that contain x relative frequency of x in the corpus [Danescu-Niculescu-Mizil et al., 2009]

Output
DE ops.

Corpus (raw text)

pNPI list

NPI contexts

S ranked

(DE learning)

(pNPI learning)

Clueless!

High-quality lists of NPIs are not available for most languages!

梏 NPIs are hard to learn from data. [Lichte and Soehn, 2007; Hoeksema 1997]
梏 Most NPIs do not translate across languages.

Where’s the orange?
S-a abținut să facă veu comentariu. [this is not Klingon!]

S_r(x) := fraction of DE contexts that contain x relative frequency of x in the corpus

Translation alternative

Not all DE operators translate.
39% of the 36 DE operators discovered in iteration 9 have no English equivalents in the largest list of English DE ops available

Does it really work for Klingon?

Q: How to choose the seed NPI?
A: Most languages have an NPI that translates to “any”.

Q: How strong is the seed?
A: It depends on the language: connection to linguistic typology. [Haspelmath, 2001]