Shallow Semantics: Semantic Role Labelling, and Beyond Shallow Semantics

Martha Palmer
University of Colorado

July 28, 2011
LING7800-006

Outline From Tuesday

- WordNet, OntoNotes Groupings, PropBank
- VerbNet
  - Verbs grouped in hierarchical classes
  - Explicitly described class properties
- FrameNet
- Links among lexical resources
  - PropBank, FrameNet, WordNet, OntoNotes groupings
- **Automatic Semantic Role Labeling with PropBank/VerbNet**

Today's Outline

- Shallow semantics: Automatic Semantic Role Labeling with PropBank/VerbNet
- Beyond shallow semantics

VerbNet: Basis in Theory

- Beth Levin, *English Verb Classes and Alternations* (1993)
- Verb class hierarchy: 3100 verbs, 47 top level classes, 193
- *"Behavior of a verb . . . is to a large extent determined by its meaning"* (p. 1)

  Amanda hacked the wood with an ax.
  Amanda hacked at the wood with an ax.
  Craig notched the wood with an ax.
  *Craig notched at the wood with an ax.
- Can we move from syntactic behavior back to semantics?

Intersective Levin Classes

- More syntactically and semantically coherent
  - sets of syntactic patterns
  - explicit semantic components
  - relations between senses

→ VERBNET

verbs.colorado.edu/~mpalmer/verbnet

Dang, Kipper & Palmer, IJCAI00, Coling00

Goals – Ex. Answering Questions

- *Similar concepts*
  - *Where are the grape arbors located?*
  - *Every path from back door to yard was covered by a grape-arbor, and every yard had fruit trees.*

VerbNet – *cover fill-9.8*

- **WordNet Senses:** ..., cover(1,2, 22, 26),..., staff(1),
- **Thematic Roles:** Agent [+animate]
  Theme [+concrete],
  Destination [+location, +region]
- **Frames with Semantic Roles**
  "The employees staffed the store"
  " The grape arbors covered every path"
  Theme V Destination

location(E,Theme,Destination)
location(E,grape_arbor,path)

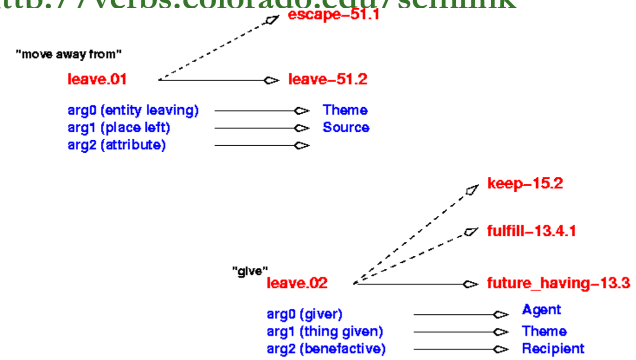
VerbNet as a useful NLP resource

- Semantic role labeling
- Inferences

*While many of the weapons used by the insurgency are leftovers from the Iran-Iraq war, Iran is still **providing** deadly weapons such as EFPs -LRB- or Explosively Formed Projectiles -RRB-.*

provide(Iran, weapons, ?Recipient) →
cause(Iran, E)
has_possession(start(E), Iran, weapons)
has_possession(end(E), ?Recipient, weapons)
transfer(during(E), weapons)

Mapping from PB to VerbNet
http://verbs.colorado.edu/semlink



Mapping from PropBank to VerbNet (similar mapping for PB-FrameNet)

Frameset id = ***ship.01***	Sense = *ship*	VerbNet class = ***Send -11.1***
Arg0	Sender	Agent/Sender*
Arg1	Package	Theme
Arg2	Recipient	Destination/ *Goal OR Recipient
Arg3	Source	Source

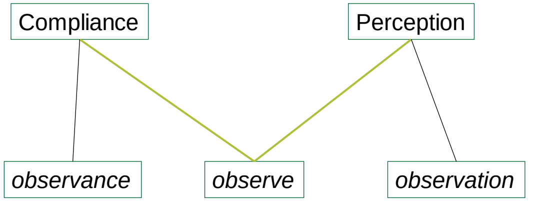
*FrameNet Labels

Baker, Fillmore, & Lowe, COLING/ACL-98
Fillmore & Baker, WordNetWKSHP, 2001

FrameNet

- Baker, Collin F., Charles J. Fillmore, and John B. Lowe. (1998) The Berkeley FrameNet project. In *Proceedings of COLING/ACL-98* , pages 86--90, Montreal.
- Fillmore, Charles J. and Collin F. Baker. (2001). Frame semantics for text understanding. In the *Proceedings of NAACL WordNet and Other Lexical Resources Workshop* Pittsburgh, June.

words, frames, lexical units



2 lexical units sharing same form:
Compliance.*observe*,
Perception.*observe*

Lexical-units: Wrap-up

Lexical units are the entities with respect to which we define
- meanings
- grammatical behavior
- semantic relations with other entities
- morphological relations with other entities

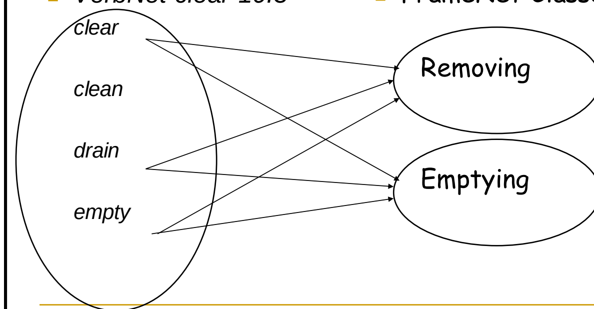
In short, there aren't interesting things to say about the verb *observe* in general, but only about the individual lexical units that happen to have the form *observe*.

Assumes lexical units can be determined

13

Mapping Issues (1)
VerbNet verbs mapped to FrameNet

- *VerbNet clear-10.3*
- FrameNet Classes



14

Mapping Issues (2)
VerbNet verbs mapped to FrameNet

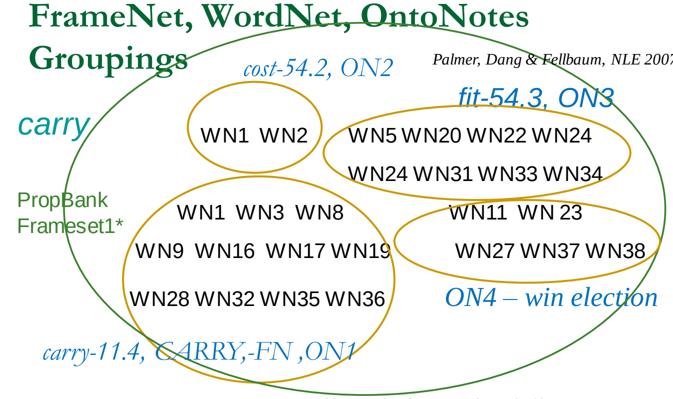
VN Class: *put 9.1*	**FrameNet frame: *place***
Members: *arrange*, immerse, lodge, mount, sling***	Frame Elements:
Thematic roles:	
• agent (+animate)	• Agent
	•Cause
• theme (+concrete)	• Theme
• destination (+loc, -region)	• Goal
Frames:	Examples:
• …	•…

*different sense
** not in FrameNet

15

SEMLINK-PropBank, VerbNet, FrameNet, WordNet, OntoNotes Groupings

Palmer, Dang & Fellbaum, NLE 2007



**ON5-ON11 carry oneself,carried away/out/off, carry to term*

SEMLINK

- Extended VerbNet 5,391 lexemes (91% PB)
- Type-type mapping PB/VN, VN/FN
- Semi-automatic mapping of PropBank instances to VerbNet classes and thematic roles, hand-corrected. (now FrameNet)
- VerbNet class tagging as automatic WSD

Brown, Dligach, Palmer, IWCS 2011

- Run SRL, map Arg2 to VerbNet roles, Brown performance improves *Yi, Loper, Palmer, NAACL07*

Automatic Labelling of Semantic Relations

- Given a constituent to be labelled
- Stochastic Model
- Features:
  - Predicate, *(verb)*
  - Phrase Type, *(NP or S-BAR)*
  - Parse Tree Path
  - Position *(Before/after predicate)*
  - Voice *(active/passive)*
  - Head Word of constituent

Gildea & Jurafsky, CL02, Gildea & Palmer, ACL02

Semantic Role Labelling Accuracy-

	FrameNet ≥ 10 inst	*PropBank*	*PropBank ≥ 10 instances*
Gold St. parses		82.8	84.1
Automatic parses	82.0	79.2	80.5

.
- *FrameNet* examples (training/test) are handpicked to be unambiguous.
- Lower performance when also deciding which constituents get labeled
- Higher performance with traces.

Progress in SRL

- Performance improved from 82.8% to 89% Colorado
  - (Gold Standard parses, < 10 instances)
  - Same features plus
    - Named Entity tags
    - Head word POS
    - For unseen verbs – backoff to automatic verb clusters
  - SVM's
    - Role or not role
    - For each likely role, for each Arg#, Arg# or not
    - No overlapping role labels allowed

Pradhan, et. al., ICDM03, Sardeneau, et. al, ACL03,
Chen & Rambow, EMNLP03,
Gildea & Hockemaier, EMNLP03

Progress in SRL

- Performance improved from 82.8% to 89% Colorado
- Penn results, original features, also SVM, 88%
  - (Gold Standard parses, < 10 instances)
  - Same features plus
    - Named Entity tags
    - Head word POS
    - For unseen verbs – backoff to automatic verb clusters
  - SVM's
    - Role or not role
    - For each likely role, for each Arg#, Arg# or not
    - No overlapping role labels allowed

Pradhan, et. al., ICDM03 & HLT-NAACL04, Sardeneau, et. al, ACL03,Chen & Rambow, EMNLP03, Gildea & Hockemaier, EMNLP03

CoNLL-04 Shared Task

Results (Gold Standard Parses)

Data	System (feature set)	P	R	F1	Class Acc
2002	G&P (Penn)	71	64	67	77.0
2002	SVM Colorado (basic)	83	79	81	87.9
2002	SVM Penn (basic)	-	-	-	93.1
2002	SVM Colorado (rich features)	89	85	87	91.0
2004	SVM Penn (basic)*	89	88	88	93.5
2004	SVM Colorado (rich features)**	90	89	89.4	93.0

**Yi and Palmer, KBCS04, ** Pradhan, et al, NAACL04*

Discussion

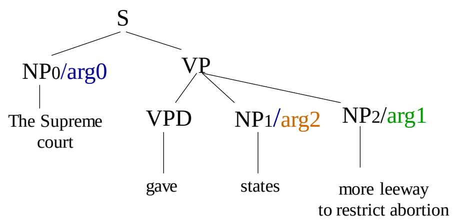
- Comparisons between Colorado and Penn
  - Both systems are SVM-based
  - **Kernel**:
    - Col: 2nd degree polynomial kernel;
    - Penn: 3rd degree kernel (radial basis function)
  - **Multi-classification**:
    - Col: one-versus-others approach;
    - Penn: pairwise approach
  - **Features**:
    - Same basic features
    - Col adds: NE, head word POS, partial path, verb classes, verb sense, head word of PP, first or last word/pos in the constituent, constituent tree distance, constituent relative features, temporal cue words, dynamic class context (Pradhan et al, 2004)
  - Kernels allow the automatic exploration of feature combinations.

Examining the classification features

- Path: the route between the constituent being classified and the predicate
- Path is not a good feature for classification
  - Doesn't discriminate constituents at the same level
  - Doesn't have full view of the subcat frame
    - doesn't distinguish subject of a transitive verb and and the subject of an intransitive verb
- Path is the best feature for identification
  - Path accurately captures the syntactic configuration between a constituent and the predicate.

Xue & Palmer, EMNLP04

Same Path – two different args



Arg1: VPD↑VP↓NP
Arg2: VPD↑VP↓NP

Possible feature combinations?

- Head word of the constituent
- POS of head word
- Phrase type
- Problem: same head word, POS, or phrase type may play different roles with regard to different verbs
- Combine with predicate

Other features

- Position + voice
- due to Colorado: Pradhan et al 2004:
  - first word of the current constituent
  - last word of the current constituent
  - left sibling of the current constituent

Results (Gold Standard Parses)

Data	System (feature set)	P	R	F1	Class Acc
2002	G&P	71	64	67	77.0
2002	SVM Colorado (basic)	83	79	81	87.9
2002	SVM Penn (basic)	-	-	-	93.1
2002	SVM Colorado (rich features)	89	85	87	91.0
2004	SVM Penn (basic)*	89	88	88	93.5
2004	SVM Colorado (rich features)**	90	89	89.4	93.0
2004	MaxEnt Penn (designated features and combinations)***	-	-	88.5	93.0

**Yi and Palmer, KBCS04, ** Pradhan, et al, NAACL04, ***Xue and Palmer, EMNLP04*

Progress on SRL

- Performance improved from 82.8% to 89% Colorado
- Penn results, different features, MaxEnt, 88%
  - (Gold Standard parses, < 10 instances)
  - Same features plus
    - Named Entity tags
    - Head word POS
    - For unseen verbs – backoff to automatic verb clusters
  - SVM's
    - Role or not role
    - For each likely role, for each Arg#, Arg# or not
    - No overlapping role labels allowed

Xue & Palmer, EMNLP -04

SRL + WSD

- Szu-ting Yi, Penn Dissertation, Chapter 8
- 2% SRL improvement with Frameset tags for 10 most highly polysemous, highly frequent verbs, ex. *call*
- Marginal improvement for verbs with > 100 instances (half +, half -)
- No improvement for verbs with < 100, > 50 instances

Portability of SRL

- Performance improved from 77% to 89%

  Automatic parses, 81% F, Brown corpus, 68%
  - Same features plus
    - Named Entity tags
    - Head word POS
    - For unseen verbs – backoff to automatic verb clusters
  - SVM's
    - Role or not role
    - For each likely role, for each Arg#, Arg# or not
    - No overlapping role labels allowed

Pradhan, et. al., ICDM03, Sardeneau, et. al, ACL03,Chen & Rambow, EMNLP03, Gildea & Hockemaier, EMNLP03, Yi & Palmer, ICON04 CoNLL-04, 05 Shared Task

Can SemLink improve Generalization?

- Overloaded Arg2-Arg5
  - PB: verb-by-verb
  - VerbNet: same thematic roles across verbs
- Example
  - Rudolph Agnew,…, was named [ARG2 {Predicate} a nonexecutive director of this British industrial conglomerate.]
  - ….the latest results appear in today's New England Journal of Medicine, a forum likely to bring new attention [ARG2 {Destination} to the problem.]
- Use VerbNet as a bridge to merge PB and FN and expand the Size and Variety of the Training

Arg1 groupings; (Total count 59710)

Group1 (53.11%)	Group2 (23.04%)	Group3 (16%)	Group4 (4.67%)	Group5 (.20%)
Theme; Theme1; Theme2; Predicate; Stimulus; Attribute	**Topic**	**Patient; Product; Patient1; Patient2**	**Agent; Actor2; Cause; Experiencer**	**Asset**

Arg2 groupings; (Total count 11068)

Group1 (43.93%)	Group2 (14.74%)	Group3 (32.13%)	Group4 (6.81%)	Group5 (2.39%)
Recipient; Destination; Location; Source; Material; Beneficiary	**Extent; Asset**	**Predicate; Attribute; Theme; Theme2; Theme1; Topic**	**Patient2; Product**	**Instrument; Actor2; Cause; Experiencer**

Process

- Retrain the SRL tagger
  - Original:
    - Arg[0-5,A,M]
  - ARG1 Grouping: (similar for Arg2)
    - Arg[0,2-5,A,M] Arg1-Group[1-6]
- Evaluation on both WSJ and Brown
- More Coarse-grained or Fine-grained?
  - more specific: data more coherent, but more sparse
  - more general: consistency across verbs even for new domains?

SRL Performance (WSJ/BROWN)

Loper, Yi, Palmer, SIGSEM07, Yi, Loper, Palmer, NAACL07

System	Precision	Recall	F-1
Arg1-Original	89.24	77.32	82.85
Arg1-Mapped	90.00	76.35	82.61
Arg2-Original	73.04	57.44	64.31
Arg2-Mapped	84.11	60.55	70.41
Arg1-Original	86.01	71.46	78.07
Arg1-Mapped	88.24	71.15	78.78
Arg2-Original	66.74	52.22	58.59
Arg2-Mapped	81.45	58.45	68.06

Limitations to Lexical Resources

- GIZA++ finds almost 80% of parallel predicate pairs in Gold Standard parallel Chinese/English PropBanks.

Alignment	GIZA++	Human Annotator
Ch.pred ↔ En.pred	53.1%	66.3%

Percentage of aligned predicates on 200 random Sentences in the Xinhua Corpus

1/3 of the predicates have no mapping in the other language.

What is meaning?

… just piling up words, one after the other, won't do much of anything until something else has been added. That something is named quite precisely by Anthony Burgess in this sentence from his novel *Enderby Outside* (1968):

- *And the words slide into the slots ordained by syntax, and glitter as with atmospheric dust with those impurities which we call meaning.*

Stanley Fish,
How to Write a Sentence: And How To Read One,p.2
From Mark Liberman's Language Log, June 14, 2011

How do we sprinkle atmospheric dust over our sentences?

- Where are we now?
- Where should we go next?
  - Inference, probably probabilistic…
- How do we get there?

Where are we now? These were some of the pieces

- We've reviewed
  - PropBanking coverage
  - Sense tagging approach
- And mentioned
  - Treebanking
  - Coreference annotation
- **Now let's put them together…**

Where we are now: Explicit Semantic Dependencies

The county coroner says he urged Saint Rita's
to move its patients.

Details of "Where we are now"

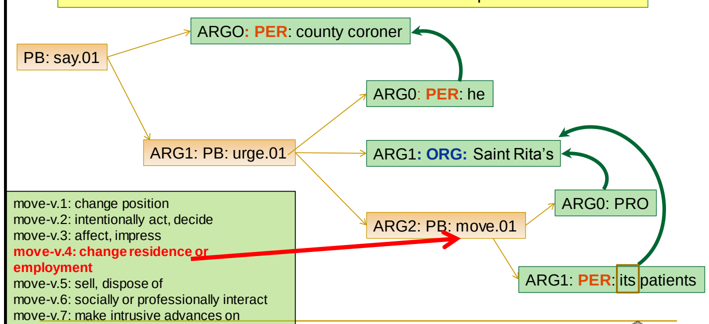
From a Broadcast Conversation story about Hurricane Katrina:

The county coroner says he urged Saint Rita's to move its patients.

```
(TOP (S (NP-SBJ (DT The)
                (NN county)
                (NN coroner))
        (VP (VBZ says)
            (SBAR (-NONE- 0)
                  (S (NP-SBJ (PRP he))
                     (VP (VBD urged)
                         (NP-1 (NNP Saint)
                               (NNP Rita)
                               (POS 's))
                         (S (NP-SBJ (-NONE- *PRO*-1))
                            (VP (TO to)
                                (VP (VB move)
                                    (NP (PRP$ its)
                                        (NNS patients)))))))))
        (. /.)))
```


Now: Explicit semantic dependencies

The county coroner says he urged Saint Rita's
PRO to move its patients.



How do we sprinkle atmospheric dust over our sentences?

- Where are we now?
- **Where should we go next?**
- How do we get there?

Where should we go next?

- *Detecting similar concepts*
- *Recovering implicit arguments*

Palmer, et. al, ACL-86, Gerber & Chai, ACL-2010

The county coroner says he urged Saint Rita's
to move its patients.

- *The eventual devastation [of Saint Rita's] threatened their lives.*
- *Did the flooding put the patients' lives in danger?*

Semantic links

- *[threaten, endanger]*
  - WordNet synsets
- *[endanger, "put in danger"]*
  - PropBank light verb construction annotation
  - Noun predicates, preposition predicates*

**Ken Litkowski working with FrameNet*

- *[cover fill9.8, location]*
  - Sense tagging, VerbNet semantic predicates

Where should we go next?

- *Detecting similar concepts*
- *Recovering implicit arguments*

Palmer, et. al, ACL-86, Gerber & Chai, ACL-2010

The county coroner says he urged Saint Rita's
to move its patients.

- *The eventual devastation [of Saint Rita's] threatened their lives.*
- *Did the flooding put the patients' lives in danger?*

Where should we go next?

- *Recovering Implicit Arguments*
  [Gerber & Chai, 2010]

*[$_{Arg0}$ The two companies] [$_{REL1}$ produce] [$_{Arg1}$ market pulp, containerboard and white paper]. The goods could be manufactured closer to customers, saving [$_{REL2}$ **shipping**] costs.*

- *Used VerbNet for subcategorization frames*
- *Coreference for implicit arguments*

VerbNet: *send-11.1* (Members: 11, Frames: 5)

includes "*ship*"

- Roles
  - Agent [+animate | +organization]
  - Theme [+concrete]
  - Source [+location]
  - Destination [+animate | [+location & -region]]
- One Frame:NP V NP PP.destination
  - example "*Nora sent the book to London.*"
  - syntax Agent V Theme {to} Destination
  - semantics motion(during(E), Theme)
    location(end(E), Theme, Destination)
    cause(Agent, E)

Argument roles for *ship*

- Agent [+animate | +organization]
- Theme [+concrete]
- Source [+location]
- Destination [+animate | [+location & -region]]

Hidden Axioms

- *[Companies] shipped [goods] to [customers].*
- THEMATIC ROLES:

AGENT V THEME SOURCE DESTINATION

- SEMANTICS
  - CAUSE(AGENT,E)
  - MOTION(DURING(E), THEME),
  - LOCATION(END(E), THEME, DESTINATION),

Hidden Axioms REVEALED!

- *[Companies] shipped [goods] to [customers].*
- THEMATIC ROLES:

AGENT V THEME SOURCE DESTINATION

- SEMANTICS
  - CAUSE(Companies, E)
  - MOTION(DURING(E), goods),
  - LOCATION(END(E), goods, customers),

Where should we go next?

- *Detecting similar concepts*
- *Recovering implicit arguments*

Palmer, et. al, ACL-86, Gerber & Chai, ACL-2010

The county coroner says he urged Saint Rita's
to move its patients.

- *The eventual devastation [of Saint Rita's] threatened their lives.*
- *Did the flooding [of Saint Rita's] put the patients' lives in danger?*

VerbNet – *move, roll 51.3.1 Class*

- **WordNet Senses:** ..., move(1,2, 3),
- **Thematic Roles:** Agent [+Intentional Control]
  Theme [+concrete],
  Location [+concrete]

  Agent V Theme
- **Frames with Semantic Roles**
  "[Saint Rita's] .... to move [its patients].

motion(during(E), ***its patients***)
cause(***Saint Rita's***, E)

Semantic links

- *[threaten, endanger]*
  - WordNet synsets
- *[endanger, "put in danger"]*
  - PropBank light verb construction annotation
  - Noun predicates, preposition predicates*

**Ken Litkowski working with FrameNet*
- *[cover fill9.8, location]*
  - Sense tagging, VerbNet semantic predicates

PropBank/VerbNet/FrameNet

- Complementary
- Redundancy is harmless, may even be useful
- PropBank provides the best training data
- VerbNet provides the clearest links between syntax and semantics
- FrameNet provides the richest semantics
- Together they give us the most comprehensive coverage
- **SEMLINK**

SEMLINK

- Extended VerbNet 5,391 lexemes (91% PB)
- Type-type mapping PB/VN, VN/FN
- Semi-automatic mapping of PropBank instances to VerbNet classes and thematic roles, hand-corrected. (now FrameNet)
- VerbNet class tagging as automatic WSD

Brown, Dligach, Palmer, IWCS 2011

- Run SRL, map Arg2 to VerbNet roles, Brown performance improves *Yi, Loper, Palmer, NAACL07*

Where should we go next?

- *Recovering implicit predicates*
- *Between Munich and LA you need less than 11 hours by plane.*
- *From Munich to LA it does not take more than 11 hours by plane.*

Where should we go next?

[Between Munich and LA $_{\text{ARGM-ADV}}$]
[you $_{\text{ARG0}}$]
[need $_{\text{REL}}$] [less than 11 hours by plane $_{\text{ARG1}}$].

- [[elided verb] [From Munich $_{\text{ARGM-DIR}}$] [to Los Angeles $_{\text{ARGM-GOL}}$] $_{\text{ARG0}}$] ,
[it] does [not $_{\text{ARGM-NEG}}$] [take $_{\text{REL-2.take10}}$]
[more than eleven hours by plane $_{\text{ARG1}}$] .

Where should we go next?

- TO FLY
[Between Munich and LA $_{\text{ARGM-ADV}}$]
[you $_{\text{ARG0}}$]
[need $_{\text{REL}}$] [less than 11 hours by plane $_{\text{ARG1}}$].
- TO FLY
[[elided verb] [From Munich $_{\text{ARGM-DIR}}$] [to Los Angeles $_{\text{ARGM-GOL}}$] $_{\text{ARG0}}$] ,
[it] does [not $_{\text{ARGM-NEG}}$] [take $_{\text{REL-2.take10}}$]
[more than eleven hours by plane $_{\text{ARG1}}$] .

Constructions allow us to

- Recognize a path prepositional phrase, and that it necessarily goes with a "MOTION" event
- If we have a MOTION event we can associate the *plane* with it as a vehicle
- Or just the *plane* itself can suggest a motion event…

Pandora's box?

- Which constructions?
- Which semantic predicates should they be associated with, give rise to?
- How to determine?
- Importance of grammatical relevance and empirical validation
- Bonial, et.al, *Incorporating Coercive Constructions into a Verb Lexicon, RELMS 2011*

How do we sprinkle atmospheric dust over our sentences?

- Where we are now?
- Where should we go next?
- **How do we get there?**

SEMLINK + Constructions + Statistics

Acknowledgments

- We gratefully acknowledge the support of the National Science Foundation Grants for , Consistent Criteria for Word Sense Disambiguation, Robust Semantic Parsing, Richer Representations for Machine Translation, and DARPA-GALE via a subcontract from BBN.
- Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

And thanks to

- Postdocs: Paul Kingsbury, Dan Gildea, Nianwen Xue,
- Students: Hoa Dang, Tom Morton, Karin Kipper Schuler, Jinying Chen, Szu-Ting Yi, Edward Loper, Susan Brown, Dmitriy Dligach, Jena Hwang, Will Corvey, Claire Bonial, Jin-ho Choi, Lee Becker, Shumin Wu
- Collaborators: Christiane Fellbaum, Anna Korhonen, Suzanne Stevenson

65

C L E A R