Thematic Roles in Linguistics

Martha Palmer
University of Colorado

LING 7800/CSCI 7000-017
September 2, 2014
Outline

- Fillmore – Cases
  - useful generalizations, fewer sense distinctions,
- Jackendoff – Lexical Conceptual Structure
  - Thematic roles are defined by the predicates they are arguments to
- Dowty – Proto-typical Agents and Patients
  - A bag of “agentive” entailments
- Levin – Verb classes based on syntax
  - syntactic behavior is a reflection of the underlying semantics
Lexical Conceptual Structures, 
Ray Jackendoff

- Decomposition into primitive semantic predicates – Thematic Relations

- Thematic roles inherit their meaning from the relations they are in
Semantic Decomposition

- Markers
  \[
  \begin{bmatrix}
  \text{HORSE} \\
  \text{RED}
  \end{bmatrix}
  \]
  \text{the red horse}

- Functions
  \[
  \text{SEE}(x,y) \quad \text{the man saw the (red) horse}
  \]
  \[
  \text{SEE}(x,\text{HORSE})
  \]
  \[
  \text{SEE}(\text{THE MAN},\text{THE HORSE})
  \]
  \[
  \text{SEE}(X1, Y1)
  \]
Five Semantic Functions

- GO
- BE
- STAY
- LET
- CAUSE
GO – Change of location

The train traveled from Detroit to Cincinatti.
The hawk flew from its nest to the ground.
An apple fell from the tree to the ground.
The coffee filtered from the funnel into the cup.

\[ \text{GO} \left( x, y, z \right) \]

THROUGH THE AIR/DOWNWARD

THEME GOES FROM SOURCE, TO GOAL
Full representation

[ event GOPOSIT
  ([thing John],
  [path FROM ([place AT (Denver)])],
  [path TO ([place AT (San Francisco)])])
  [MANNER: Drivingly]]
Satellite framed vs. Verb framed motion verbs – basis of LCS Interlingua

Verb-framed: French, Spanish

GO (Theme, Source, Goal)

Manner

Traverse the lake by swimming

Satellite-framed: English

GO (Theme, Source, Goal)

Manner

Swim across the lake.
Mapping from Syntax to Semantics

/flaj/
+ V
+ [NP¹_____ (from NP²) (to NP³)]
  GO (NP¹,NP²,NP³)
  THROUGH THE AIR
Max is in Africa.
The vine clung to the wall.
The dog is on the left of the cat.
The circle contains/surrounds the dot?

BE\((x,y)\)
THEME IS AT LOCATION

BE (THE DOG, LEFT OF (THE CAT))
The bacteria stayed in his body.
Stanley remained in Africa.
Bill kept the book on the shelf.

\[ \text{STAY}(x, y) \]
\[ \text{THEME IS AT LOCATION for a duration} \]

\[ \text{STAY (STANLEY, AFRICA) (for two years)} \]
The train traveled from Detroit to Cincinatti.

Harry gave the book to the library.

The book belonged to the library.
Locational modes: POSIT, POSS, ID

*The bacteria stayed in his body.*

\[
\text{STAY} \ (x, z) \\
\text{POSIT}
\]

*The library kept the book.*

\[
\text{STAY} \ (x, z) \\
\text{POSS}
\]
Locational modes: POSIT, POSS, ID

*The coach changed from a handsome young man to a pumpkin.

\[ \text{GO}_{\text{IDENT}} (x,y,z) \]

Princess Mia changed from an ugly duckling into a swan.

\[ \text{GO}_{\text{IDENT}} (x,y,z) \]

Universal grammar?
Causation and Permission

CAUSE and LET

*The rock fell from the roof to the ground.*

\[[\text{GO}_{\text{POSIT}} (x,y,z)]\]

*Linda lowered the rock from the roof to the ground.*

\[[\text{CAUSE} (a, \text{GO}_{\text{POSIT}} (x,y,z))]\]

*Linda dropped the rock from the roof to the ground.*

\[[\text{LET} (a, \text{GO}_{\text{POSIT}} (x,y,z))]\]
Linda lowered the rock from the roof to the ground with a cable.

\[
\text{CAUSE } (a, \text{ GO}_{\text{POSIT}} (x, y, z))
\]

Inst: \(i\)

Instruments only occur with causation.

CAUSE always has an \textit{event} second argument.

Dollie caused Martin to be happy.
## Lexical Conceptual Structure

<table>
<thead>
<tr>
<th>concept</th>
<th>POSIT</th>
<th>POSS</th>
<th>IDENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO</td>
<td>go</td>
<td>receive</td>
<td>become</td>
</tr>
<tr>
<td>motional</td>
<td>fall</td>
<td>inherit</td>
<td>change</td>
</tr>
<tr>
<td>BE</td>
<td>be</td>
<td>have</td>
<td>be</td>
</tr>
<tr>
<td>punctual</td>
<td>contain</td>
<td>own</td>
<td>seem</td>
</tr>
<tr>
<td>STAY</td>
<td>stay</td>
<td>keep</td>
<td>stay</td>
</tr>
<tr>
<td>durational</td>
<td>remain</td>
<td>remain</td>
<td>remain</td>
</tr>
<tr>
<td>CAUSE(a,GO)</td>
<td>bring, take</td>
<td>obtain, give</td>
<td>make, elect</td>
</tr>
<tr>
<td>CAUSE(a,STAY)</td>
<td>keep, hold</td>
<td>keep, retain</td>
<td>keep</td>
</tr>
<tr>
<td>LET(a,GO)</td>
<td>drop, release</td>
<td>accept, fritter</td>
<td>leave</td>
</tr>
<tr>
<td>LET(a,BE)</td>
<td>leave, allow</td>
<td>permit</td>
<td></td>
</tr>
</tbody>
</table>
Rules of inference

\[ \text{CAUSE}(a, \text{event}) \rightarrow \text{event}. \]
Issues

- Ducks vs. Geese?
- Abstract concepts?
Thematic Proto-Roles and Argument Selection, *David Dowty*

- Role definitions have to be determined verb by verb, and with respect to the other roles

- Thanks to Michael Mulyar for slides
Context of Dowty’s work

- Thematic relations
  - (Gruber 1965, Jackendoff 1972)

- Traditional thematic roles types include:
  - Agent, Patient, Goal, Source, Theme, Experiencer, Instrument

- “Argument-Indexing View”: thematic roles objects at syntax-semantics interface, determining a syntactic derivation or the linking relations.
Problems with Thematic Role Types

- Fragmentation: Cruse (1973) subdivides Agent into four types.
- Ambiguity: Andrews (1985) is Extent, an adjunct or a core argument?
- Symmetric static predicates: e.g. “This is similar to that” Distinct roles or not?
- Searching for a Generalization: What is a Thematic Role?
Proto-Roles

- Event-dependent Proto-roles introduced
- Prototypes based on shared entailments
- Grammatical relations such as subject related to observed (empirical) classification of participants
- Typology of grammatical relations
- Proto-Agent
- Proto-Patient
Proto-Agent

Properties

- Volitional involvement in event or state
- Sentience (and/or perception)
- Causing an event or change of state in another participant
- Movement (relative to position of another participant)
- (exists independently of event named)
  *may be discourse pragmatic
Proto-Patient

Properties:

- Undergoes change of state
- Incremental theme
- Causally affected by another participant
- Stationary relative to movement of another participant
- (does not exist independently of the event, or at all) *may be discourse pragmatic
Argument Selection Principle

- For 2 or 3 place predicates
- Based on empirical count (total of entailments for each role).
  - Greatest number of Proto-Agent entailments → Subject;
  - greatest number of Proto-Patient entailments → Direct Object.
- Alternation predicted if number of entailments for each role similar (non-discreteness).
Worked Example: Psychological Predicates

Examples:

<table>
<thead>
<tr>
<th>Experiencer Subject</th>
<th>Stimulus Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$ likes $y$</td>
<td>$y$ pleases $x$</td>
</tr>
<tr>
<td>$x$ fears $y$</td>
<td>$y$ frightens $x$</td>
</tr>
</tbody>
</table>

Describes “almost the same” relation

Experiencer: sentient (P-Agent)
Stimulus: causes emotional reaction (P-Agent)

Number of proto-entailments same; but for stimulus subject verbs, experiencer also undergoes change of state (P-Patient) and is therefore lexicalized as the patient.
Diathesis Alternations

Alternations:
- Spray / Load
- Hit / Break

Non-alternating:
- Swat / Dash
- Fill / Cover
Spray / Load Alternation

Example:

Mary loaded the hay onto the truck.
Mary loaded the truck with hay.

Mary sprayed the paint onto the wall.
Mary sprayed the wall with paint.

- Analyzed via proto-roles, not e.g. as a theme / location alternation.
- Direct object analyzed as an Incremental Theme, i.e. either of two non-subject arguments qualifies as incremental theme. This accounts for alternating behavior.
Hit / Break Alternation

John hit the fence with a stick.
John hit the stick against a fence.

John broke the fence with a stick.
John broke the stick against the fence.

- Radical change in meaning associated with break but not hit.
- Explained via proto-roles (change of state for direct object with break class).
Fill / Cover are non-alternating:

* Bill filled the tank (with water).
* Bill covered the ground (with a tarpaulin).

Only goal lexicalizes as incremental theme (direct object).
Dowty’s Hierarchy (English)

AGENT > INSTRUMENT
     > BENEFACTIVE

PATIENT > SOURCE
          > GOAL
Conclusion

- Dowty argues for Proto-Roles based on linguistic and cognitive observations.
- Three main areas of analysis: symmetric predicates, diathesis alternations, unaccusativity
- Objections: Are P-roles empirical (hit class)? Are P-roles event dependent (possibly in need of revision, e.g. something like p-patients named by event vs. p-patients defined by event)?
Motivation: From Sentences to Propositions

Who did what to whom, when, where and how?

Powell met Zhu Rongji

Powell and Zhu Rongji met

Powell met with Zhu Rongji

Powell and Zhu Rongji had a meeting

Proposition: meet(Powell, Zhu Rongji)

meet(Somebody1, Somebody2)

When Powell met Zhu Rongji on Thursday they discussed the return of the spy plane.

meet(Powell, Zhu) discuss([Powell, Zhu], return(X, plane))
Analysts have been expecting a GM-Jaguar pact that would give the U.S. car maker an eventual 30% stake in the British company.
Analysts have been expecting a GM-Jaguar pact that would give the US car maker an eventual 30% stake in the British company.
PropBank roles – based on Dowty

- PropBank Frame for *break*:

  Frameset **break.01** “break, cause to not be whole”:
  - Arg0: breaker
  - Arg1: thing broken
  - Arg2: instrument
  - Arg3: pieces

Why numbered arguments?

- Lack of consensus concerning semantic role labels
- Numbers correspond to verb-specific labels
- Arg0 – Proto-Agent, and Arg1 – Proto-Patient, (Dowty, 1991)
- Args 2-5 are highly variable and overloaded – poor performance
Consistent argument labels across different syntactic realizations

- Uuuuuusually...
  - Arg0 = agent, experiencer
  - Arg1 = patient, theme
  - Arg2 = benefactive / instrument / attribute / end state
  - Arg3 = start point / benefactive / instrument / attribute
  - Arg4 = end point
Function tags for modifiers

- Variety of ArgM’s:
  - **TMP** - when? *yesterday, 5pm on Saturday, recently*
  - **LOC** - where? *in the living room, on the newspaper*
  - **DIR** - where to/from? *down, from Antartica*
  - **MNR** - how? *quickly, with much enthusiasm*
  - **PRP/CAU** - why? *because … , so that …*
  - **REC** - himself, themselves, each other
  - **GOL** - end point of motion, transfer verbs? *To the floor, to Judy*
  - **ADV** - hodge-podge, miscellaneous, “nothing-fits!”
  - **PRD** - this argument refers to or modifies another: *…ate the meat raw*
Statistics also revealed that Taiwanese business investment is tending to increase.
Statistics also revealed that Taiwanese business investment is tending to increase.
Statistics also revealed that Taiwanese business investment is tending to increase.
Statistics also revealed that Taiwanese business investment is tending to increase.
Why do we need Frameset ID’s?

PropBank Frames Files: tend.01, care for

Roles:

Arg0: tender
Arg1: thing tended (to)

Example: John tends to the needs of his patrons.
Arg0: John
REL: tend
Arg1: the needs of his patrons
Sense distinctions in PropBank – coarse-grained

PropBank - Frames Files: tend.02, have a tendency

Roles:
Arg1: Theme
Arg2: Attribute

Example: *The cost, or premium, tends to get fat in times of crisis.*

Arg1: *The cost, or premium*
REL: *tend*
Arg2: *to get fat in times of crisis.*
Statistics also revealed that Taiwanese business investment is tending to increase.
Actual data for *leave*

Leave .01 “move away from” Arg0 rel Arg1 Arg3
Leave .02 “give” Arg0 rel Arg1 Arg2

*sub-ARG0 obj-ARG1* 44
*sub-ARG0* 20
*sub-ARG0 NP-ARG1-with obj-ARG2* 17
*sub-ARG0 sub-ARG2 ADJP-ARG3-PRD* 10
*sub-ARG0 sub-ARG1 ADJP-ARG3-PRD* 6
*sub-ARG0 sub-ARG1 VP-ARG3-PRD* 5
*NP-ARG1-with obj-ARG2* 4
*obj-ARG1* 3
*sub-ARG0 sub-ARG2 VP-ARG3-PRD* 3