Compositionality and the Theory of Argument Selection

James Pustejovsky

LOT Winter School, Nijmegen

January 8-12, 2007

COURSE OUTLINE

- Monday: Variations in Argument Selection
  What is Compositionality?
- Tuesday: Generative Lexicon as a Theory of Selection
  Types and Compositional Mechanisms
- Wednesday: Derivations using GL's Type
  Composition Logic
- Thursday: Inherent and Selectional Polysemy
  Applications of Generative Selection
- Friday: Verbal Polymorphisms
  Event Structure and Selection

Questions Addressed in this Course

- Encoding Context through Lexical Typing
- Architecture: How does the grammar encode or license selection in language?
- Expressiveness: How much argument information can and should be selected/selectable?
- Descriptiveness: How does selection contribute to coverage of the data?
- Qualia as Types: Can qualia structure and event structure be adequately modeled as types directly?

Lecture 1. Framing the Problem

- Variations in Argument Selection
- Methodology of Empirically-driven Semantics
- Defining Compositionality

The Role of Annotation in Linguistic Theory

- Semantic annotation is critical for robust language understanding:

  Summarization, question answering, inference

- Annotation schemata should focus on a single coherent theme:

  Different linguistic phenomena should be annotated separately over the same corpus

- Annotations must be consistent with each other:

  Unification and merging of multiple annotation is necessary

- The Annotate, Train, and Test Model advances linguistic theory:

  Theories needs testing to evaluate coverage and predictive force.

Hence

Semantic theories are too complex to develop without this model.

Examples of Semantic Annotation

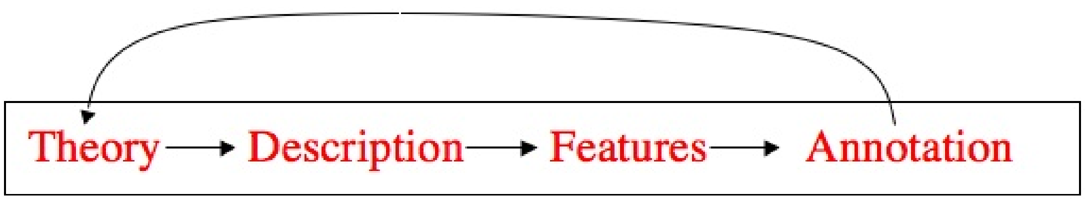
- Predicators and their named arguments:
  [The man]$_{agent}$ painted[the wall]$_{patient}$.
- Anaphors and their antecedents:
  [The protein] inhibits growth in yeast. [It] blocks production . . .
- Acronyms and their long forms:
  [Platelet-derived growth factor] (known as [pdgf]) impacts . . .
- Semantic Typing of entities:
  [The man]$_{human}$ fired [the gun]$_{firearm}$

Selectional Information encodes Compositional Structure

- [The man]$_{agent}$ painted[the wall]$_{patient}$.
  V(agent,patient):
  V(human,surface):
- [The man]$_{human}$ fired [the gun]$_{firearm}$
  V(human,firearm):
  V(agent,patient):

Methodology of Empirically-Grounded Semantics

- Annotation scheme: assumes a given feature set.
- Feature set: encodes specific structural descriptions and properties of the input data.
- Structural descriptions: theoretically-informed attributes derived from empirical observations over the data.



The Model-Annotate-Test Paradigm

Assumptions

- Language meaning is compositional.
- Compositionality is a desirable property of a semantic model.
- Many linguistic phenomena appear non-compositional.
- Generative Lexicon exploits richer representations and rules to fix holes in the compositionality model.
- Richer representations involve Lexical Decomposition.
- Richer rules involve Coercion, Subselection, Co-composition.

- Mechanisms of Selection
  - Strong Selection: Selection of arguments through typing
  - Weak Selection: Patterns of use, stereotypical collocates
- Encoding event structure in the type language
- The Principles of Decompositionality

The Principle of Compositionality

The meaning of a complex expression is determined by its structure and the meanings of its constituents.

The Principle of Compositionality

The meaning of a complex expression is determined by its structure and the meanings of its constituents.

Questions . . .

The Principle of Compositionality

The meaning of a complex expression is determined by its structure and the meanings of its constituents.

Questions . . .

What is the nature of the structure?

What is the meaning of a constituent?

What counts as a constituent?

Challenges to Simple Compositionality

(1) a. Mary began [to eat her breakfast].
  b. Mary began [eating her breakfast].
  c. Mary began [her breakfast].

(2) a. Mary enjoyed her beer.
  b. John enjoys his coffee in the morning.
  c. Bill enjoyed the movie.

Verbal Polysemy

Fillmore (1985), Levin and Rappaport (1998), Jackendoff (1990), Pustejovsky (1995)

(3) a. The woman baked a potato in the oven.
    b. The woman baked a cake in the oven.

(4) a. John swept.
    b. John swept the floor.
    c. John swept the dirt into the corner.
    d. John swept the dirt off the sidewalk.
    e. John swept the floor clean.
    f. John swept the dirt into a pile.

shovel, rake, shave, weed.

(5) a. John whistled.
b. John whistled at the dog.
c. John whistled a tune.
d. John whistled a warning.
e. John whistled her appreciation.
f. John whistled to the dog to come.

yell, snap, whisper.

Unaccusatives and Selectional Specificity

(6) Externally Caused Events: break, etc.

a. The vase broke.

b. Mary broke the vase.

c. The storm broke the window.

(7) Internally Caused Events: decay, bloom, etc.

a. The flowers bloomed early.

b. *The gardener bloomed the flowers.

Adjectival Selection

Adjectives modify specific aspects of the head Noun:

(8) a. a former professor
b. my former car
c. the retired admiral

(9) a. the escaped prisoner
b. an unbaked cake

(10) a. the vacation/well-built house
b. a fast/young typist
c. some bright/expensive bulbs
d. a long/scratched record/CD.

(11) a. a blue pen

b. the stone lion

c. the toy gun

Most Composition is Considered Function Application

1. What is the nature of the function?
2. What does it apply to; i.e., what can be an argument?

1. John loves Mary.
2. love(Arg$_1$,Arg$_2$)
3. Apply love(Arg$_1$,Arg$_2$) to Mary
4. $\Longrightarrow$ love(Arg$_1$,Mary)
5. Apply love(Arg$_1$,Mary) to John
6. $\Longrightarrow$ love(John,Mary)

Entity Types and Predicate Types

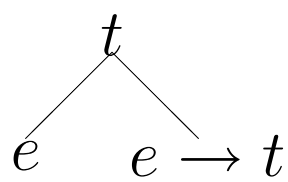
The Simply Typed λ-Calculus

(a) e is a type.

(b) t is a type.

(c) If a and b are types, then $a \longrightarrow b$ is a type.

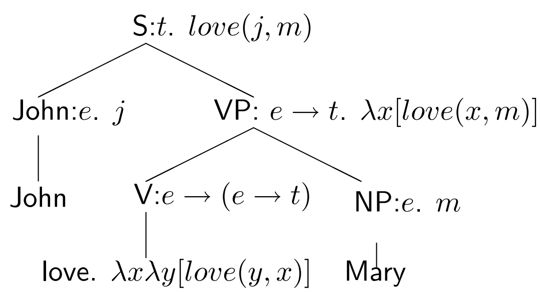
A simple type tree:



Function Application: If α is of type a, and β is of type $a \longrightarrow b$, then $\beta(\alpha)$ is of type b.

Compositionality in Practice

(a) John loves Mary.



The Notion of Selection in a Compositional Theory

1. What elements can select?
2. What is an argument?
3. What does it mean for a predicate to select an argument?
4. How does selection relate to composition and lexical decomposition?

Verb Meaning: The Predicative Complex

1. Properties of the participants
2. Change (of being, state, location, relation)
3. Causation and means
4. Manner of the activity
5. Temporal and Locational constraints
6. Intentionality of the actor
7. Instrumental and co-agentive information
8. Psychological state of the participants

Property Abstraction

(12) a. What kinds of properties are abstracted as selectional restrictions?
b. Which of these aspects of meaning can be abstracted into arguments of the predicate?

Predicative Structure

(1) a. Verb: V

b. Arguments: x, y, z, ...

Predicative Structure

(1) a. Verb: V How do we decompose the meaning?
b. Arguments: x, y, z, ...

Predicative Structure

(1) a. Verb: V How do we decompose the meaning?
    b. Arguments: x, y, z, ...

(2) a. Body: the predicate, with bound variables.
    b. Binding Environment (Args): the parameter list.

Predicative Structure

(1) a. Verb: V How do we decompose the meaning?
    b. Arguments: x, y, z, ...

(2) a. Body: the predicate.
    b. Binding Environment (Args): the parameter list.

$$\overbrace{\lambda x_i}^{Args}\overbrace{[\Phi]}^{Body}$$

Decomposition Strategies

1. ATOMIC PREDICATION:
   Parameter structure mirrors syntactic behavior:

$$\lambda x_n \ldots \lambda x_1[\Phi]$$

2. PARAMETRIC DECOMPOSITION:
   Parameter structure adds additional arguments for interpretation in the model, where $m > n$:

$$\lambda x_m \ldots \lambda x_{n+1} \lambda x_n \ldots \lambda x_1[\Phi]$$

Decomposition Strategies

3. SIMPLE PREDICATIVE DECOMPOSITION:
   P is defined as a complex expression of subpredicates over the parameter:

$$\lambda x[\Phi_1, \ldots \Phi_k]$$

4. FULL PREDICATIVE DECOMPOSITION:
   Parameter structure is enhanced, and P is defined as a complex of subpredicates:

$$\lambda x_m \ldots \lambda x_{n+1} \lambda x_n \ldots \lambda x_1 [\Phi_1, \ldots \Phi_k]$$

5. SUPRALEXICAL DECOMPOSITION:
Parameter structure is enriched through mechanism of additional operators, while P is enriched by a complex compositional operation:

$$\lambda f_\sigma \lambda x_1[\mathcal{R}(f)(x_1)](\lambda x[\Phi_1, \ldots \Phi_k])_\sigma$$

Atomic Predication

(13) Parameter structure mirrors the syntactic behavior:

$$\text{Verb}(\text{Arg}_1, \ldots, \text{Arg}_n) \implies$$

$$\lambda x_n \ldots \lambda x_1[\Phi]$$

(14) a. $\lambda x[\text{die}(x)]$

b. The flower died.

(15) a. $\lambda y \lambda x[\text{hit}(x, y)]$

b. The car hit the wall.

Parametric Decomposition

(16) Parameter structure adds additional arguments for interpretation in the model, where $m > n$:

$$\text{Verb}(\text{Arg}_1, \ldots, \text{Arg}_n) \implies$$

$$\lambda x_m \ldots \lambda x_{n+1} \lambda x_n \ldots \lambda x_1 [\Phi]$$

(17) a. $\lambda y \lambda x \lambda e[\text{kill}(e, x, y)]$: (Davidson, 1967)
b. The gardner killed the flower.

(18) a. $\lambda z \lambda y \lambda x \lambda e[\text{go}(e, x, y, z)]$: (Hobbs, 1993)
b. Nicholas went to China.

Simple Predicative Decomposition

(19) P is defined as a complex expression of subpredicates over the parameter:

$$\text{Verb}(\text{Arg}_1) \implies$$

$$\lambda x[\Phi_1, \ldots \Phi_k]$$

(20) a. **die**: $\lambda x[\text{alive}(x) \wedge \text{Become}(\neg\text{alive}(x))]$

b. The flower died.

(21) **bachelor**:

$\lambda x[\text{male}(x) \wedge \text{person}(x) \wedge \text{adult}(x) \wedge \neg\text{married}(x)]$

Full Predicative Decomposition

(22) Parameter structure is enhanced, and P is defined as a complex of subpredicates:

$$\text{Verb}(\text{Arg}_1, \ldots, \text{Arg}_n) \quad \Longrightarrow$$

$$\lambda x_m \ldots \lambda x_{n+1} \lambda x_n \ldots \lambda x_1 [\Phi_1, \ldots \Phi_k]$$

(23) a. **kill**:
$\lambda y, x, e_1, e_2[\text{act}(e_1, x, y) \wedge \neg\text{dead}(e_1, y) \wedge \text{dead}(e_2, x) \wedge e_1 < e_2]$:

Supralexical Decomposition

$$\mathsf{Verb}(\mathsf{Arg}_1, \ldots, \mathsf{Arg}_n) \quad \Longrightarrow \lambda x_n \ldots \lambda x_1[\Phi]$$

$$v \Longrightarrow \lambda f_\sigma \lambda x_1[\mathcal{R}(f)(x_1)]$$

$$\Longrightarrow \lambda f_\sigma \lambda x_1[\mathcal{R}(f)(x_1)](\lambda x[\Phi])_\sigma$$

$$\Longrightarrow \lambda x_1[\mathcal{R}([\Phi])(x_1)]$$

How is Compositionality related to Selection?

- The parameters that we give a relation are values of the tuple that satisfies the conditions under which this relation is judged true;
- They are also determinants on the members of this tuple.

How Arguments are Interpreted

(24) a. A rock fell.

$\exists x \exists e[\mathsf{fall}(e, x) \wedge \mathsf{rock}(x)]$

b. A rock died.

$\exists x \exists e[\mathsf{die}(e, x) \wedge \mathsf{rock}(x)]$

Predicate Abstraction

Consider the computation involved in interpreting these sentences:

(25) a. The woman slept soundly.
    b. The soldier died in the street.
    c. The child dreamt of Christmas.

The “Fail Early” Strategy of Selection

(26) Arguments can be viewed as pretests for performing the action in the predicate.

(27) If the argument conditions are not satisfied, the predicate does not get interpreted.

Argument Typing as Abstracting from the Predicate

(28) Richer typing for arguments:
i. Identifies specific predicates in the body of the expression that are characteristic functions of an argument;
ii. pulls this subset of predicates out of the body, and creates a $pretest$ to the expression as a restricted quantification over a domain of sorts, denoted by that set of predicates.

Argument Typing as Abstracting from the Predicate

(29)

$$\lambda x_2 \lambda x_1 [\Phi_1, \ldots \overbrace{\Phi_{x_1}}^{\tau}, \ldots \overbrace{\Phi_{x_2}}^{\sigma}, \ldots, \Phi_k]$$

σ and τ are sets of predicates describing properties of arguments to the predicate complex.

Predicate Abstractions Become Argument Types

(30)

$$\lambda x_2 : \sigma \ \lambda x_1 : \tau[\Phi_1, \ldots, \Phi_k - \{\Phi_{x_1}, \Phi_{x_2}\}]$$

(31) σ and τ have now become reified as types on the arguments.

Reified Types Determine Syntactic Selection

Count/Mass

(32) a. {not much/all/lots of} gold/water/dirt/sand
b. {every/two/several} chairs/girls/beaches

Arity Constraints

(33) a. Mary arrived.
b. John hit Mary.
c. Mary gave a book to John.

Animacy

(34)a. The man / the rock fell.

b. The man / *the rock died.

(35)a. John forced / *convinced the door to open.

b. John forced / convinced the guest to leave.

Spatial Constraints

(36)a. John poured milk into / * on his coffee.

b. John poured water into / on the bowl.

(37)a. John crossed the lake / *the ladder.

b. John climbed *the ocean / the building.

Selecting for Manner

(38)a. The children behaved themselves.

b. Mary behaved *(well).

(39)a. John performed *(admirably).

b. John performed his piece yesterday.

Collections and Number

(40)a. The crowd / *the man assembled.

b. The audience / the fans dispersed into the streets.

(41)a. The professor counted his students.

b. He had ten graduates and five undergraduates.

c. He had only one.

count(human,[number])

Generative Lexicon is a Typed Decomposition Formalism

(42) a. A Full Parametric Decomposition Theory.

b. Employs the computational efficiency of pretest argument typing.

c. Qualia Structure provides a restricted calculus over decomposition potential .

GL's Theoretical Starting Points

- The human conceptual apparatus (i.e. the ability to categorize and represent the world) is one of generative categorization and compositional thought (as opposed to extensional).
- The human linguistic capacity reflects our ability to categorize and represent the world in the particular ways we do.
- Therefore, language is a natural manifestation of our generative construction of the world through the categories it employs.

GL Addresses the Open Texture of Language

- The infinite variability of reference in language (Waismann, 1951)
- A direct product of the essential incompleteness of terms and their composition.
- building a notion of open texture directly into word meaning, the formal mechanisms which give rise to sentence meanings ensure that both analytic and contextual aspects of meaning are available.

Merging Two Traditions in Study of Language

- Corpus Language Philosophy:
  Manipulation of usage situations associated with words and word tuples.
- Generative Lexicon:
  Encoding lexical dynamic context for richer interpretation of natural language.

Generative Lexicon

- Two classes of sortal constraints on a concept:
  - Argument structure
  - Event structure
- These bind into the Qualia Structure
- Compositional Rules invoke
  - Type Selection: Exact match of the type
  - Type Accommodation:The type is inherited
  - Type Coercion: Type selected must be satisfied

Argument and Body in Generative Lexicon

(43)

$$\overbrace{\underbrace{\lambda x_n \ldots \lambda x_1}_{}^{AS}\;\overbrace{\lambda e_m \ldots \lambda e_1}^{ES}}^{Environ}\quad \overbrace{[Q_1 \wedge Q_2 \wedge Q_3 \wedge Q_4;\ C]}^{Body}$$

AS: Argument Structure

ES: Event Structure

Q_i: Qualia Structure

C: Constraints

Qualia Structure

Formal: the basic category which distinguishes it within a larger domain;

Constitutive: the relation between an object and its constituent parts;

Telic: its purpose and function;

Agentive: factors involved in its origin or “bringing it about”.

GL Logical Form

(44)

$$\begin{bmatrix} \alpha \\ \text{ARGSTR} = \begin{bmatrix} \text{ARG1} = x \\ \ldots \end{bmatrix} \\ \text{EVENTSTR} = \begin{bmatrix} \text{EVENT1} = e1 \\ \text{EVENT2} = e2 \end{bmatrix} \\ \text{QUALIA} = \begin{bmatrix} \text{CONST} = \textbf{what } x \textbf{ is made of} \\ \text{FORMAL} = \textbf{what } x \textbf{ is} \\ \text{TELIC} = e_2\textbf{: function of } x \\ \text{AGENTIVE} = e_1\textbf{: how } x \textbf{ came into being} \end{bmatrix} \end{bmatrix}$$

Compositionality

- Local context modeled as Strong Selection.
- Compositional rules refer to these types:
- Types can be selected;
- Types can be accommodated.
- Types can be exploited (coercion).
- Types can be introduced (coercion).
- Composition can license new interpretations (cocompose).

Mechanisms of Selection refer to Type Level

Types of Expressions in Language (Level)

- Natural Types: atomic concepts of FORMAL, CONST, and AGENTIVE;
- Artifactual Types: Adds concepts of TELIC;
- Complex Types: Cartesian types formed from both Natural and Artifactual types.

Level of Types in the Major Categories

1. Noun

   N: rock, water, woman, tiger, tree

   F: knife, beer, husband, dancer

   C: book, lunch, university, temperature

2. Verb

   N: fall, walk, rain, put, have

   F: donate, spoil, quench

   C: read, perform

3. Adjective

N: red, large, flat

F: useful, good, effective

C: rising

Mechanisms of Selection

- Pure Selection: The type a function requires is directly satisfied by the argument.
- Accommodation: The type a function requires is inherited by the argument.
- Coercion: The type a function requires is wrapped around the argument, embedding it within the required type.

GL Lexical Structure

(45)

$$
\begin{bmatrix}
\alpha \\
\text{ARGSTR} = \begin{bmatrix} \text{ARG1} = x \\ \ldots \end{bmatrix} \\
\text{EVENTSTR} = \begin{bmatrix} \text{EVENT1} = e1 \\ \text{EVENT2} = e2 \end{bmatrix} \\
\text{QUALIA} = \begin{bmatrix} \text{CONST} = \textbf{what } x \textbf{ is made of} \\ \text{FORMAL} = \textbf{what } x \textbf{ is} \\ \text{TELIC} = e_2\textbf{: function of } x \\ \text{AGENTIVE} = e_1\textbf{: how } x \textbf{ came into being} \end{bmatrix}
\end{bmatrix}
$$

Motivating the Notion of Natural Kind

(46) a. Nominal Predication: How the common noun behaves predicatively;

b. Adjectival Predication: How adjectives modifying the the common noun can be interpreted;

c. Interpretation in Coercive Contexts: How NPs with the common noun are interpreted in coercive environments.

(47) a. Mary saw every dog/pet.

b. John visited a $man/doctor$.

c. $Birds/planes$ can fly.

(48)a. a sick dog/pet

b. an American $man/doctor$

c. white $birds/planes$

(49)a. Otis is a dog.

b. Otis is a poodle.

b. Eno is a cat.

(50)a. ?Otis is a dog and and an animal.

b. !That is a dog and a cat.

c. Otis is a dog and therefore an animal.

(51)a. !This box is large and small.

b. !Your gift is round and square.

(52) a. This is both a pen and a knife.

b. The substance is a stimulant and an anti-inflammatory.

(53) a. Mary is a housewife and a doctor.

b. Bernstein was a composer and a conductor.

(54) a. This object is a knife and therefore a weapon.

b. Emanuel Ax is a pianist and therefore a musician.

(55) Emanuel Ax is a pianist and therefore a human.

(56) a. very old gold

b. a new tree
c. a young tiger
d. such a beautiful flower

(57) a. a blue/Swiss pen
b. a bright/expensive bulb
c. a long/shiny CD

(58) a. a very old friend
b. a good professor
c. such a beautiful dancer

(59) a. Mary enjoyed drinking her beer.
b. Mary enjoyed her beer.

(60) a. John began to write his thesis.

b. John began writing his thesis.
c. John began his thesis.

(61) a. !John finished the tree.
b. !Mary began a tiger.

Lecture 2. Theory Meets Corpus

- Enriching Composition Operations for More Explanatory Coverage
- Corpus Data on Semantic Transformations
- Lexical Sets and Corpus Pattern Analysis

Recall the Themes of this Course

- Language meaning is constructed compositionally
- Semantic theory constructs models over artificial data
- Compositionality must account for corpus data
- Corpus data must be annotated with existing theories
- Corpus phenomena force revisions and enrichments of theory.

What have we Learned thus far?

- Principle of Compositionality and Functional Typing
- The type of a verb is derived from the types of it arguments
- Argument selection is argument typing

Compositionality in Language

- One phenomenon we haven't looked at yet. . .

Compositionality in Language

- One phenomenon we haven't looked at yet. . .
- Logical Polysemy

Logical Polysemy

- When a single word or phrase has the ability to appear in selected contexts that are contradictory in type specification.

(62)a. Mary doesn't believe the book.

b. John sold his book to Mary.

a. believe: V(human,proposition)

Apply believe(human,proposition) to [the book]

b. sell: V(human,physobj,human)

Apply sell(human,physobj,human) to [the book]

Types in Generative Lexicon: Pustejovsky (2001,2007), Asher and Pustejovsky (2005)

The Type Language

(63)a. e the type of entities; t, truth values.

(σ, τ range over simple types, and subtypes of e.)

b. If σ and τ are types, then so is $\sigma \rightarrow \tau$;

c. If σ and τ are types, then so is $\sigma \bullet \tau$;

d. If σ and τ are types, then so is $\sigma \otimes_Q \tau$, for $Q =$ $const(c)$, $telic(t)$, or $agentive(a)$.

Qualia Structure as Types

(64) TYPE FEATURE STRUCTURE:

$$\begin{bmatrix} x : \alpha \\ \text{QUALIA} = \begin{bmatrix} \text{CONST} : & \beta \\ \text{FORMAL} : & \alpha \\ \text{TELIC} : & \tau \\ \text{AGENTIVE} : & \sigma \end{bmatrix} \end{bmatrix}$$

Qualia as Types

(65)
$$\begin{bmatrix} x : & & \alpha \\ & \bigotimes_c & \beta \\ & \bigotimes_t & \tau \\ & \bigotimes_a & \sigma \end{bmatrix}$$

Natural Entities

Entities formed from the application of the FORMAL and/or CONST qualia roles:

For the predicates below, e_N is structured as a join semi-lattice, $\langle e_N, \sqsubseteq \rangle$;

(66)a. *physical*, *human*, *stick*, *lion*, *pebble*

b. *water*, *sky*, *rock*

Natural Entity Types as a Lattice

(67)