Lecture 10: Computational Language Acquisition

Afra Alishahi
January 19, 2009
Human Language Acquisition

• **Representation** of the linguistic knowledge
  • What is innate, what is learnable?
  • How is the knowledge organized in mind and brain?
    • Are there separate areas/levels for representing lexical/syntactic/semantic knowledge?

• **Acquisition** of the linguistic knowledge
  • What are the processes involved in language learning?
  • Are different types of knowledge acquired in order?
Learnability in Acquisition

• Representation of the linguistic knowledge
  • What is innate, what is learnable?
  • How is the knowledge organized in mind and brain?
    • Are there separate collections/levels for representing lexical/syntactic/semantic knowledge?

• Acquisition of the linguistic knowledge
  • What are the processes involved in language learning?
  • Are different types of knowledge acquired in order?
Modularity in Acquisition

- Representation of the linguistic knowledge
  - What is innate, what is learnable?
  - How is the knowledge organized in mind and brain?
    - Are there separate collections/levels for representing lexical/syntactic/semantic knowledge?

- Acquisition of the linguistic knowledge
  - What are the processes involved in language learning?
  - Are different types of knowledge acquired in order?
Syntax vs. Semantics

• How is the **surface structure** (i.e., syntax) linked to the **underlying meaning** (i.e., semantics)?
  
  • Alternative 1: syntax is learned independently of semantics
  
  • Alternative 2: syntax and semantics are learned at the same time

• A suitable case study: **verb argument structure**

• The relationship between the semantics of verbs and their syntactic form
Verb Argument Structure

- Knowledge of verb argument structure:
  - Number and type of arguments that the verb takes
    - *The man built the house
    - *The man built, *The house built the man
  - Semantic roles that verb arguments receive in an event
    - She<Agent> broke the window<Theme>, The window<Theme> broke
    - *She<Agent> broke
  - Syntactic realization of the verb and its arguments
    - I filled the glass with water
    - *I filled water into the glass
General Regularities

- Challenges of argument structure acquisition:
  - Detecting general regularities: young children are aware of a general mapping between syntactic forms and semantic elements
    
    * bunny gorped duck ⇒ causal action? 
    * kitty blicked down the street ⇒ manner of motion?

- Learning idiosyncrasies: highly similar verbs (e.g. load, fill, pour) have different syntactic behaviour
  
  * I filled the glass with water, *I filled water into the glass
  * They loaded the truck with hay, They loaded hay into the truck
Mapping of Syntax to Semantics

- Semantic Bootstrapping (Pinker, 1984)

- The syntactic behaviour of a verb is innately determined by the decompositional representation of its meaning

- *Agent* is 1st argument of *CAUSE*, *Theme* is 1st argument of *GO* & *BE*, *Patient* is 2nd argument of *CAUSE*.

- With the innate knowledge of the mapping between semantics and syntax, a child can predict the correct mapping once she knows what a verb means

- It fails to explain recent experimental findings.
Experimental Findings

- **Item-based learning:**
  - Young children build their linguistic knowledge around individual items
  - Two year olds show little tendency to apply syntactic structures they have already learned to new verbs
  - Tendency to generalize familiar constructions to new forms increases as children grow older

→ Verb-Island Hypothesis (Tomasello, 1992)
U-shaped Learning Curve

- Observed **U-shaped** learning curves in children

- **Imitation**: an early phase of conservative language use (each verb is used in the constructions it has been seen in before)

- **Generalization**: knowledge of general regularities is acquired and applied to new forms

- **Overgeneralization**: occasional mis-application of general patterns, which leads to errors

- **Recovery**: over time, overgeneralization errors cease to appear in child speech
Akhtar’s (1999) experiment

- 2-4 years old children were taught novel verbs in non-standard English word order (e.g., SOV)

  Look! Bunny duck gorped.

- In elicited production, 2 and 3-year-olds imitated the observed pattern half the time and “corrected” to the English SVO pattern half the time.

- 4-year-olds rarely imitated the observed order, almost always correcting to the SVO order
Computational Generalization

- Computational models of learning general regularities from input data
- Emergence of abstract knowledge from exposure to a number of instances (McClelland & Kawamoto, 1986; Allen 1997)
- Allen (1997):
Overgeneralization & Recovery

- Overgeneralization errors happen in different domains of language
  - English past-tense: *I goed*
  - Argument structure: *You can drink me the milk*

- Consistent patterns among children:
  - For a given form, errors are few at the early steps
  - Number of errors increase as general patterns emerge
  - After a while errors decrease again
Lack of Negative Evidence

- **Negative evidence**
  - Information about which strings of words do not belong to language (corrective feedback from parents)
- Marcus (1993): there is no reliable negative evidence available to children
  - Some suggest that, even if corrective feedback is provided, children ignore it.
- Recovery from overgeneralization must occur **without** relying on negative evidence
Recovery Mechanisms

- Many learning mechanisms are suggested as factors in recovery from overgeneralization (Goldberg, 1999; MacWhinney, 2004)
  - Entrenchment, competition, cue construction, ...

- Recently, probabilistic interaction between various factors is suggested as a solution (Onnis et al., 2002; Alishahi & Stevenson, 2008)
  - The frequencies of verbs and general constructions
  - Semantic match between a construction and an event
Productive Generalization

• Children eventually stop overgeneralizing, but productive use of language continues through adulthood:

  The truck rumbled down the hill.

  The fly buzzed into the room.

• Alternative: **Construction Grammar** (Lakoff 1987, Fillmore et al. 1988, Langacker 1999)

• In addition to the idiosyncratic meanings associated with individual words or morphemes, meaning may also be *directly* associated with syntactic forms
Construction Grammar

- Argument structure construction (Goldberg, 1995)
- A mapping between underlying verb-argument relations and the syntax used to express them

\[ Sub \ j \ V \ Ob \ j \ Ob \ j2 \leftrightarrow X \ cause \ Y \ receive \ Z \]

Example: *Pat faxed Bill the letter.*

\[ Sub \ j \ V \ Oblique \leftrightarrow X \ move \ Y \]

Example: *The fly buzzed into the room.*
How are Constructions Learned?

- **Tomasello (1991):**
  - Argument structure patterns are initially acquired on a verb-by-verb basis
  - Constructions associated with the common syntactic patterns are learned through a process of categorization and generalization over the input.

- **Goldberg (1995):**
  - Constructional meaning is formed around the meanings of highly frequent *light verbs*
  - E.g., the construction “Subj V Obl” paired with the meaning “X moves Y” corresponds to the light verb *go*
Computational Models of Constructions

- FrameNet (Baker, Fillmore, Low, 1998): a database of lexical constructions (or frames)
- The acquisition of constructions
  - Learning lexical constructions (Chang, 2004)
  - Learning verb meaning from image data (Dominey, 2003; Dominey & Inui, 2004)
  - Learning abstract constructions from verb usage data (Alishahi & Stevenson, 2008)
Chang (2004)

• A model for learning lexical-based multi-word constructions from child-directed data

• **Goal**: learn associations between form and meaning relations

• **Learning task**: finding the best grammar to fit the observed data
Alishahi & Stevenson (2008)

• A Bayesian, usage-based model of early argument structure acquisition
  • Each verb usage is viewed as a set of features
  • Constructions are viewed as a probability distribution over syntactic and semantic features
  • A Bayesian clustering method detects and groups similar usages to form constructions
Verb usages as argument structure frames:

*Sara is eating an apple*

<table>
<thead>
<tr>
<th>head verb</th>
<th>eat</th>
</tr>
</thead>
<tbody>
<tr>
<td>verb semantic primitives</td>
<td>[act,consume]</td>
</tr>
<tr>
<td>argument roles</td>
<td>&lt;Agent, Theme&gt;</td>
</tr>
<tr>
<td>argument categories</td>
<td>&lt;human, food&gt;</td>
</tr>
<tr>
<td>syntactic pattern</td>
<td>arg1 verb arg2</td>
</tr>
</tbody>
</table>
• Constructions as clusters of similar frames:

- **Alishahi & Stevenson (2008)**

**Syntactic pattern:**

**Argument categories:**

**Verb semantic primitives:**
Constructions as clusters of similar frames:
Verb Semantic Roles

- Semantic (thematic) roles, such as Agent, Theme and Instrument, indicate the relations of the participants in an event to the main predicate.

*Pat gave the hammer to Matt.*

\[\text{Give}_{\text{cause,possess}}(\text{Pat, Hammer, To(Matt)})\]
Main Questions

- What is the nature of semantic roles?
  - **Traditional view**: roles are atomic and universal, such as Agent, Theme, Goal, … (e.g., Jackendoff 1990)
  - **Proto-role Hypothesis** (Dowty, 1991): roles are a set of properties, such as volitional, affecting, animate

- Where do they come from?
  - **Traditional view**: roles and their link to syntactic positions are innate (e.g., Pinker 1989)
  - **Alternative view**: they are gradually learned from verb usages (e.g., Tomasello 2000)
Learnability of Thematic Roles

- Usage-based account: verb-specific roles change to general roles over time

- Experimental evidence confirms that access to general roles such as Agent and Theme is age-dependent (Shayan & Gershkoff-Stow, 2007)
Children are sensitive to the association between semantic roles (e.g., Agent) and grammatical functions (e.g., Subject) from an early age.

Fisher 1994, 1996; Nation et al., 2003

Nativist account: innate “linking rules” that map roles to sentence structure enable children to infer associations between role properties and syntactic positions (e.g., Pinker, 1989).
Computational Studies of Roles

• Assignment of general pre-defined roles to sentence constituents
  • E.g., McClelland and Kawamoto (1986), Allen (1997)

• Role learning
  • Learning verb-specific roles from annotated data (Chang 2004)
  • Discovering relational concepts from unstructured examples (Kemp et al., 2006; Doumas et al., 2008)
  • Acquiring semantic profiles for general roles from verb usages (Alishahi & Stevenson, 2008)
Open Questions

• How various aspects of language acquisition interact with each other?

• Various learning procedures are most likely interleaved (e.g., word leaning and syntax acquisition)

• Most of the existing models of language acquisition focus on one aspect, and simplify the problem

• How to evaluate the models on realistic data?

• Large collections of child-directed utterances/speech are available, but no such collection of semantic input