A model of frame and verb compliance in language acquisition

Rutvik Desai
Computer Science and Cognitive Science
Indiana University
Word learning

- Children learn new words rapidly
- Observe the object or event and associate it with the word
  - Not enough
- Use another source of information, syntax
  - *Gleitman (1990)*
    - Correlate syntactic structure with events, objects, and words
Frame and Verb Compliance

- One way to study the effect of syntax on the acquisition of word meaning is to place familiar words in unfamiliar or incorrect syntactic context
  - Transitive verbs into intransitive frames
  - Intransitive verbs into transitive frames

- If the interpretation is in accordance with the frame, children are said to be *Frame Compliant*; otherwise they are *Verb Compliant*. 
Frame and Verb Compliance

- Subjects: 2.5 to 12 year old children; adults
- Task: Enact scenes described in sentences using wooden toy animals as props
  - *The lion puts in the ark
  - *The zebra brings
  - *The elephant comes the giraffe

  Naigles, Fowler, and Helm (1992); Naigles, Gleitman, and Gleitman (1993)

- Results
  - Younger children, especially the 2-year-olds, were Frame Compliant. They changed the meaning of the verb to fit the sentence.
  - Older children, especially the adults, were Verb Compliant.
  - Intermediate ages, intermediate levels of compliance
Overgeneralization

• Children’s early verb use is mostly correct
  \textit{Tomasello (2000)}

• A major exception to this occurs somewhere around age 3
  \textit{Bowerman (1977; 1982)}

• They sometimes use verbs in incorrect frames
  – *Don’t fall that on me

• The must eventually learn which uses are licensed
  – \textit{sink} can be used either transitively or intransitively
  – \textit{fall} and \textit{go} allow only intransitive, noncausal interpretation

• How to eliminate overgeneralizations?
Theories of Compliance/Overgeneralization
Maturation

• Verbs get organized in *narrow range subclasses* based on semantic criteria
• If one verb alternates (between transitive and intransitive use), other verbs in the same subclass are assumed to alternate too (Frame Compliance)
• Those subclasses for which there has been no evidence of alternation are “closed” at maturation, and no new information is accepted (Verb Compliance)

*Pinker (1989)*

• There are factors other than age that affect compliance behavior
• No across-the-board shift from frame to verb compliance
  – Some verbs elicit more compliance than others; some frames too
  – Timing of the shift varies for different verbs and frames
Theories of Compliance/Overgeneralization

Mutual Exclusivity

- Children will allow only one lexical entry to occupy a semantic niche
- If two words are determined to have similar meanings, one of them is preempted and removed from the lexicon
  - Causative (over-generalized) come is similar to bring
  - When causative come is used, bring is non-existent
  - When bring becomes more frequent, causative come declines

*Bowerman (1982); Clark (1987, 1991); Markman (1987)*

- While ME may have some role to play, it does not account for all the data
  - Why some verbs elicit Verb Compliance earlier than others
  - For some verbs, it is difficult to find similar meaning verbs that can preempt their use in the right way
Theories of Compliance/Overgeneralization
Lexical Knowledge

- Children’s conjunctures about verb meanings are refined by ongoing events as well as by the structures in which they appear.
- At early stages, open-minded children assume that not all structures have yet been heard and therefore certain properties (such as causality) may be unknown (Frame Compliance).
- Older children and adults feel warranted to believe that all relevant information about the meaning has been obtained; a novel structure is ill formed (Verb Compliance).

*Naigles, Gleitman, and Gleitman (1993)*

- Explains many effects in the data well.
- Amount of compliance varies for some verbs depending on the frame.
- How exactly do children determine if they have enough knowledge of the meaning?
Theories of Compliance/Overgeneralization
Lexical Knowledge and Innate Principles

• “The learning device is asking itself, in effect: Assuming Principles [the Theta Criterion and the Projection Principle], what could be the meaning of the verb now heard, such that these principles projected this observed (surface) structure for it? Such a deductive procedure will be invoked only when the learner does not have secure knowledge of the verb in question.” (p. 37)

*Lidz, Gleitman, and Gleitman (2001)*

• Amount of compliance varies for some verbs depending on the frame
• How exactly do children determine if they have “secure” knowledge of the meaning?
The Model

- Task: comprehension
- Input: utterances generated by a grammar, one word at a time
- Output: Representation of a “scene” described by the sentence
  - One or two objects and an action
  - Predict entire scene after each word
  - The representation of an action includes four units encoding causality (1100 ⇒ causal event; 0011 ⇒ noncausal event)
Network Architecture
Grammar

S → NP | NP1 | NP is IV | NP1 are IV | NP is TV NP
NP → DET N | DET SIZE N
NP1 → NP and NP
N → boy | girl | dog | mouse | ...
SIZE → large | small
IV → jumping | dancing | running | ...
TV → pushing | holding | hugging | ...
DET → a

• a boy (N)
• a small dog and a girl (NN)
• a boy is jumping (NV)
• a mouse and a large frog are jumping (NNV)
• a small boy is pushing a large dog (NVN)
Comprehension

- Utterances probabilistically divided into training and testing set
- Differing probabilities for different sentence types based on frequency; about 1/3 in the training set overall
- Close to 100% accuracy on the training set, about 96% on the test set
Modeling Compliance

- The network trained in stages with increasing vocabulary.
- Two transitive and two intransitive verbs constant all stages.
- The number of nouns increased gradually; trained to 100% accuracy at each stage.
- Tested at each stage using the utterances containing the words seen at that stage for generalization.
- Test for compliance: insert incompatible verbs in transitive and intransitive frames with novel nouns.
  - *a xxx is dancing a yyy*

- Define a “causality index” \( \delta \):
  - \( \delta = \) mean activation of the first two units – mean activation of the last two units.
  - Positive \( \delta \) ⇒ causal/transitive; negative \( \delta \) ⇒ noncausal/intransitive interpretation.
δ across various stages
• Early stages
  – Rote-learning
  – “Item-based” or context-bound representations
  – Predicts poor generalization
  – Stronger form-form correlations
• Later stages
  – Categorical, context-free representations
  – Predicts good generalization
  – Stronger form-meaning correlations
• What is nature of representations in the network?

• What is nature of representations in children?
Generalization across various stages

- The performance on the training set at each stage is ~100%

ICDL 2004
Principal Components

- The utterances used at the first stage are passed through the network, and hidden unit activations recorded after each word at each stage.

- If the network is using a simple strategy (such as the presence or absence of a single feature) to recognize utterances, there would be few underlying dimensions of variation; few principal components required.

- For more complex processing, the representations would change gradually as new information is taken into account as the utterance is processed; more principal components required.
The values of Principal Components at various stages
Representations in children

- Many recent studies suggest that much of children's early competence is rote learned or item-based
  - Organized around concrete linguistic items, phrases, words, not based on system-wide syntactic categories and schemas
    - *Tomasello (2000)*
  - Lieven et al. (1997): 12 children, 2-3 years age
    - Virtually all verbs and predicative terms were used in only 1 sentence frame
    - 92% of all utterances were from 25 lexically based patterns (different for different children)
  - Pine and Lieven (1997)
    - When children started to use ‘a’ and ‘the’ they did it for almost totally different set of nouns
    - No abstract category for the determiner
Representations in children

• Tomasello (1992): Verb islands
  • Some verbs were used in only one sentence frame (Cut ___) some were used in more complex frames (Draw___, Draw___on___, etc.)
  • Morphological markings were uneven across verbs
  • The best predictor of a child's use of a given verb was not her use of other verbs, but the use the same verb preceding days
  • No transfer of structure across verbs

• Dabrowska (2000):

<table>
<thead>
<tr>
<th>Age</th>
<th>%Formulaic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1;6-1;11</td>
<td>94.8</td>
</tr>
<tr>
<td>2;0-2;5</td>
<td>91.2</td>
</tr>
<tr>
<td>2;6-2;11</td>
<td>77.6</td>
</tr>
<tr>
<td>3;0-3;9</td>
<td>63.5</td>
</tr>
</tbody>
</table>
Cross-linguistic evidence

- Italian (Pizutto & Caselli, 1994): 47% of verbs were used in only one frame, another 40% were used in 2 or 3 frames.

- Portuguese (Brazilian) (Rubino & Pine, 1998): Verb use corresponded to high-freq. patterns in the input.

- Hebrew (Berman, & Armon-Lotem, 1995) First 20 verbs were rote-learned and morphologically unanalyzed.

- Also, Hungarian, German, Dutch, Spanish, Russian, Inuktitut, Catalan.
What about the input?

- The shift from Frame to Verb compliance in the network depends on the changing input
  - Increasingly combinatorial input with increasing vocabulary

- Is this, in any way, similar to what children experience? Isn’t changing the input to your model bad?
The nature of input to children

• First, even if children live in a constant linguistic environment, with time they necessarily sample more and more utterances from it
  – More words, more word combinations experienced with time
• Second, the linguistic environment of children changes with time
  – When talking to young children, the words used by caregivers are less diverse
  – Child-directed speech (CDS) is syntactically and semantically simplified
  – The type-token ratio in CDS decreases with age
  – Caregivers are sensitive to the child’s comprehension level; not so much to their production

  *Cameron-Faulkner and Tomasello (2003); Chapman (1981); Cross 1977; Hu (1994); Pine (1994)*
Summary and conclusions

- The shift in the model’s behavior emerges from a shift in the underlying representations, which in turn is caused by a change in the input
  - Initially, context-bound, item-based, rote-learned; components undifferentiated
  - Later, context-free, categorical, or referential; individual components separated
Summary and conclusions

• No explicit decision making, no specialized principles

• “open-mindedness” in early stages and “secure knowledge” of the meaning in later stages can be explained as relative strengths of form/form and form/meaning correlations

• Shift in compliance emerges from the nature of input, the task, and characteristics of the network

• Specialized-looking behaviors such as Frame and Verb compliance can be understood as cases of pattern completion, a ubiquitous phenomenon
When pattern completion is not used

Wait! Wait! …Cancel that, I guess it says ‘helf’.