Generalizations are driven by semantics and constrained by statistical preemption

New evidence from artificial language experiments

Florent Perek & Adele Goldberg

University of Birmingham & Princeton University
• Learning a language = generalizing beyond the input

• For instance, using verbs in novel ways

\textit{It meeked} (witnessed form)


• Overgeneralization errors \cite{Bowerman1990}

?? \textit{Don’t giggle me}
## Generalizing beyond the input

- When and why do speakers generalize beyond their input? And when and why do they not?

- What aspects of the input are relevant?
  - Does language learning *only* consist of gleaning statistical regularities in the input?
  - What about the role of the *function* of constructions?

### Generalizations

<table>
<thead>
<tr>
<th>Previous work</th>
<th>Exp. 1: Generalization</th>
<th>Exp. 2: Preemption</th>
<th>Conclusion</th>
</tr>
</thead>
</table>


Artificial language learning studies
(e.g., Casenhiser & Goldberg 2004; Finley & Badecker 2009; Folia et al. 2010; Fedzechkina et al. 2010; Hudson Kam & Newport 2005; Wonnacott et al. 2008)

- Participants exposed to novel <utterance, video scene> pairs
- Statistical structure of input is manipulated
- To test the role of statistics in language learning
Experiment 1

- Two word order constructions: APV and PAV, a suffix –po on the patient argument
  
  the panda the pig-po mooped (APV: Agent Patient-po Verb)
  the pig-po the panda mooped (PAV: Patient-po Agent Verb)
  ‘the panda mooped the pig’

- Six novel verbs (e.g., glim, moop, wub) referring to transitive actions (e.g., ‘punch’, ‘push’, ‘head-butt’)

- Two test conditions
  
  - Lexicalist: 3 APV-only verbs, 3 PAV-only verbs
  - Alternating: 2 APV-only, 2 PAV-only, 2 alternating verbs
Constructions are rarely synonymous in natural languages (cf. Bolinger 1968; Givon 1979; Goldberg 1995)

Our two constructions differ in the intensity of the effect on the patient

- APV: strong effect: the patient rapidly moves across the screen and out of the scene with dramatic gestures
- PAV: weak effect: the patient hardly moves, with similar but less ample gestures
Example of APV exposure pair

the monkey the panda-\textit{po} glimmed
<table>
<thead>
<tr>
<th>Generalizations</th>
<th>Previous work</th>
<th>Exp.1: Generalization</th>
<th>Exp.2: Preemption</th>
<th>Conclusion</th>
</tr>
</thead>
</table>

**Example of PAV exposure pair**

the panda-*po* the monkey glimmed
• Participants: 24 Princeton undergraduates (18-22, 16 female)

• Exposure (2 days)
  – 36 sentence-scene pairs, each verb used 6 times
  – Participants were asked to repeat each sentence

• Sentence production task
  – Participants described new scenes; verb was given
  – Each of the 6 verbs presented 4x, twice each with video showing strong and weak effect
  – Two new novel verbs, not witnessed in the input
Example of production trial (strong effect)

what happened here?
(pilked)
the pig the cat-\textit{po} pilked

\textit{or}
the cat-\textit{po} the pig pilked
Example of production trial (weak effect)

what happened here?
(pilked)
the pig the cat-\textit{po} pilked
\textit{or}
the cat-\textit{po} the pig pilked
To what extent do speakers generalize constructions to unattested verbs?

- Hypothetical data: conservative, verb-based behavior

<table>
<thead>
<tr>
<th>Generalizations</th>
<th>Previous work</th>
<th>Exp.1: Generalization</th>
<th>Exp.2: Preemption</th>
<th>Conclusion</th>
</tr>
</thead>
</table>

![Graph: APV production vs. PAV production](image)

APV-only verbs in input

PAV-only verbs in input
To what extent do speakers generalize constructions to unattested verbs?

- Hypothetical data: full generalization across verbs
Experiment 1: Results

Alternating condition: two alternating verbs

Lexicalist condition: no alternating verbs

Full generalization
Mixed effects logistic regression
(to predict the probability of producing APV)

- Strong tendency to produce APV when the effect is strong ($\beta = 3.4756$, $p < 0.0001$)

- APV-only verbs tend to be used (slightly) more often with APV compared to novel verbs ($\beta = 0.8111$, $p = 0.0013$)

- Interaction between Condition and Effect: the effect of the functional difference is weaker in the lexicalist condition ($\beta = -1.1113$, $p = 0.0085$)
Summary of Experiment 1

• Tendency for participants to generalize (using verbs in the contextually appropriate constructions)

• They may ignore usage of individual verbs

• Linguistic function can overcome statistical information in the choice of construction

• Contrasts with Wonnacott et al.’s (2008) results with synonymous constructions; see also Perek & Goldberg (2015); Thothathiri & Rattinger (2016)
<table>
<thead>
<tr>
<th>Generalizations</th>
<th>Previous work</th>
<th>Exp.1: Generalization</th>
<th>Exp.2: Preemption</th>
<th>Conclusion</th>
</tr>
</thead>
</table>

**Discussion**

- The meaning of constructions is a source of productivity in natural language (e.g., Goldberg 1995)

- But constructional generalizations are typically restricted, e.g.,

  *Explain me this.* \(\rightarrow\) *(Explain this to me)*

- **Statistical preemption:** (Goldberg 1995; Goldberg 2006, Boyd & Goldberg 2011; Robenalt & Goldberg 2015, 2016)

  Repeated occurrence of a form A when a different form B is expected provides evidence that only A is acceptable
Experiment 2: statistical preemption

• Similar design to Experiment 1

• 1 PAV-only verb statistically preempted from APV
  i.e., used with both strong and weak effect in PAV in exposure

• Will speakers only use the verb in PAV contexts, regardless of strength of effect?

• Will this affect the way they learn the language?
Experiment 2: Results

![Bar chart showing results for SOV and OSV verb use in Strong and Weak effect conditions.]

Novel verbs tend to be used with the contextually appropriate construction.
Experiment 2: Results

Preempted PAV verb tends to be used in PAV

Novel verbs tend to be used with the contextually appropriate construction
Experiment 2: Results

**APV-only verbs**
- Strong: Used more than PAV
- Weak: Used less than PAV

**PAV-only verbs**
- Strong: Used more than APV
- Weak: Used less than APV

**Preempted PAV verb**
- Preempted by APV

**Novel verbs**
- Used with contextually appropriate construction

**Generalizations**
- Previous work
  - Exp.1: Generalization
  - Exp.2: Preemption
- Conclusion
Mixed effects logistic regression
(to predict the probability of producing APV)

- Again, tendency to produce APV when the effect is strong ($\beta = 2.0433, p < 0.0001$)

- But mitigated by strong effects of VerbType: participants are more conservative with all verbs
  - APV-only: $\beta = 1.3727, p = 0.0002$
  - PAV-only: $\beta = -1.2858, p = 0.0013$
  - preempted PAV: $\beta = -1.4558, p = 0.0026$
Summary of Experiment 2

• Productions with the new novel verbs show speakers did learn the functional difference between constructions.

• Speakers are also very sensitive to preemptive information; they used it to infer the restriction on the preempted verb.

• They were also more lexically conservative with other verbs:

  Unlike Exp. 1, APV-only and PAV-only verbs were mostly used with APV and PAV, respectively.

• Preemptive exposure for one verb provides evidence that other verbs, too, are restricted in their distributions.
Conclusion

• Adult learners are sensitive to the form and function of newly learned constructions

  Speakers are willing to generalize beyond their input according to the function of constructions

• They are also sensitive to the distribution of verbs

  Statistical preemption provides evidence that verbs are restricted in their distributions
<table>
<thead>
<tr>
<th>Generalizations</th>
<th>Previous work</th>
<th>Exp.1: Generalization</th>
<th>Exp.2: Preemption</th>
<th>Conclusion</th>
</tr>
</thead>
</table>

Thanks for your attention!

f.b.perek@bham.ac.uk       adele@princeton.edu


