Broad generalizations about pragmatic processing

Christina Kim
University of Chicago
Linguistics

Universality and Empirical Validity/LSA workshop
July 12, 2013
How do we generalize at the right level of abstraction?

overgeneralization ↔ overfitting*

*“The central problem of linguistics is overfitting” –JRiggle (the internet)
How do we generalize at the right level of abstraction?

overgeneralization $\leftrightarrow$ overfitting

> Often data is sparse.

> Data is noisy, subject to selection bias, other non-interesting variability.

*“The central problem of linguistics is overfitting” –JRiggle (the internet)
How do we generalize at the right level of abstraction?

overgeneralization $\leftrightarrow$ overfitting*

> Often data is sparse.
> Data is noisy, subject to selection bias, other non-interesting variability.
> We want to make broad generalizations (more universal = better).

*“The central problem of linguistics is overfitting” –JRiggle (the internet)
How do we generalize at the right level of abstraction?

overgeneralization ↔ overfitting*

> Often data is sparse.
> Data is noisy, subject to selection bias, other non-interesting variability.
> We want to make broad generalizations (more universal = better).
> Linguistic data is really intricate/complex, so we gotta be careful.

*“The central problem of linguistics is overfitting” –JRiggle (the internet)
overgeneralization ↔ overfitting

> Some form-based element conveys some (family of) meanings universally (e.g. fronting = focus; fronting = topic; scrambling = specificity; rising intonation = uncertainty; etc.)
getting at Language rather than language

overgeneralization ↔ overfitting

> Some form-based element conveys some (family of) meanings universally (e.g. fronting = focus; fronting = topic; scrambling = specificity; rising intonation = uncertainty; etc.)

> Characterization of a broad phenomenon in terms of narrow linguistic properties s.t. you can’t generalize to a language with different superficial properties (e.g. focus → \{prosodic prominence, displacement, morphological marking\}).
where Language meets the world:
where Language meets the world:

> pragmatic processes are systematic, “rule-governed” (not outside of the grammar/“linguistic competence”)

Chris Kim  Broad generalizations about pragmatic processing
where Language meets the world:

> pragmatic processes are systematic, “rule-governed” (not outside of the grammar/“linguistic competence”)

> BUT: the information content involved in pragmatic processes include extra-linguistic knowledge
where Language meets the world:

> pragmatic processes are systematic, “rule-governed” (not outside of the grammar/“linguistic competence”)

> BUT: the information content involved in pragmatic processes include extra-linguistic knowledge

> our knowledge about types of situations in the world, likelihoods of certain outcomes relative to others given a set of circumstances
where Language meets the world:

> pragmatic processes are systematic, “rule-governed” (not outside of the grammar/“linguistic competence”)

> BUT: the information content involved in pragmatic processes include extra-linguistic knowledge

> our knowledge about types of situations in the world, likelihoods of certain outcomes relative to others given a set of circumstances

> the mental representation of the concept advisor and its conceptual associates for \{people who have gone through grad school, financial planners, civilians\}

linguistic and non-linguistic knowledge comes from (an individual’s) cumulative experience with the world
roadmap

1 Narrow and broad explanations of the same data.
   
   **Case study 1.** Category effects in the processing of alternative-sensitive particles

2 Making fast generalizations accurately.
   
   **Case study 2.** Imprecision and cues to goal structure (a.k.a. Have you ever played flip-cup?)
roadmap

1. **Narrow and broad explanations of the same data.**

   **Case study 1.** Category effects in the processing of alternative-sensitive particles

2. **Making fast generalizations accurately.**

   **Case study 2.** Imprecision and cues to goal structure (a.k.a. Have you ever played flip-cup?)

Chris Kim

Broad generalizations about pragmatic processing
how do we infer the right focus alternatives given a context of use?

Veronica only met LUCY.

**Context-invariant part:**
Veronica met Lucy.
Veronica didn’t meet any other person in $A(C)$.

**Context-dependent part:**
$A(C) =$ the salient group of people such that Veronica didn’t meet anyone in that group (except Lucy).

Rooth [1985,1992]
Veronica only met LUCY.

Context-invariant part:
Veronica met Lucy.
Veronica didn’t meet any other person in $A(C)$.

Context-dependent part:
$A(C) =$ the salient group of people such that Veronica didn’t meet anyone in that group (except Lucy).
$= \{\text{Bobby’s ex-girlfriends}\}$

Rooth [1985,1992]
Strategy

**What we want:**
A model of the context as it is represented in the minds of language users.
What we want:
A model of the context as it is represented in the minds of language users.

Strategy:
Use comprehenders’ expectations about likely discourse continuations to learn about their mental representations of the context, how different information types are encoded.
Linking expectations to a behavioral measure

The Visual World paradigm
(Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy [1995])

- **Logic:** Based on prior experience, we have expectations about how a situation will resolve given what has already occurred.

- **Applied to language:** Based on prior linguistic experience, we have expectations about how a sentence/discourse will resolve, given the prior context.
Focus alternatives: conceptual category

context sentence:
1a. Lauren has some apples and some oranges.
1b. Lauren has some pears and some oranges.
1c. Lauren has some boots and some sandals.

target sentence:
2a. Nathan has some apples.
2b. Nathan only has some apples.
Focus alternatives: conceptual category

**context sentence:**
1a. Lauren has *some apples* and *some oranges*.
1b. Lauren has *some pears* and *some oranges*.
1c. Lauren has *some boots* and *some sandals*.

**target sentence:**
2a. Nathan has *some apples*.
2b. Nathan only has *some apples*.

← [−Only]
← [+Only]
Focus alternatives: conceptual category

**context sentence:**
1a. Lauren has some apples and some oranges. ← [+Mention]
1b. Lauren has some pears and some oranges.
1c. Lauren has some boots and some sandals.

**target sentence:**
2a. Nathan has some apples. ← [−Only]
2b. Nathan only has some apples. ← [+Only]
Focus alternatives: conceptual category

context sentence:
1a. Lauren has some apples and some oranges. ← [+Mention]
1b. Lauren has some pears and some oranges. ← [+Novel/+Category]
1c. Lauren has some boots and some sandals. ← [+Novel/−Category]

target sentence:
2a. Nathan has some apples. ← [−Only]
2b. Nathan only has some apples. ← [+Only]
Focus alternatives: conceptual category

average point of disambiguation

Time (ms) from target word onset

[−Only]  [+Only]

Chris Kim  Broad generalizations about pragmatic processing
Focus alternatives: conceptual category

average point of disambiguation

[+Novel/−Category]  [+Novel/+Category]  [+Mention]

Time (ms) from target word onset

[−Only]  [+Only]

Chris Kim
Broad generalizations about pragmatic processing
Focus alts are constrained by recent discourse mention.

1 **theory1**: It’s all pretty low-level.

Conceptual associates get activated during lexical access, resulting in similarity-based facilitation (No additional explanation necessary.)
Focus alts are constrained by recent discourse mention.

1. **theory1**: It’s all pretty low-level.
   Conceptual associates get activated during lexical access, resulting in similarity-based facilitation (No additional explanation necessary.)

2. **theory2**: It’s top-down.
   We’re constantly reasoning about/making inferences about our interlocutors’ mental states, beliefs, desires; THAT’S fundamentally what determines what’s salient.
predictions/implications

1 theory1: conceptual representations, the organization of the mental lexicon

>‘Lexical’ categories — activated bottom-up, with closer associates more activated than distant ones (Collins & Loftus [1975]; Tversky & Hemenway [1984]; Cree, McRae & McNorgan [1999])

>Strength of category facilitation should be directly related to similarity/distance in conceptual space
theory 2: comprehension involves understanding communicative goals/speaker intent

> Ad hoc categories, categories based on world knowledge — rule out continuations that are unlikely given a real world situation (Barsalou [1982,1983])

> Expect different category effects depending on whether there is a salient goal, whether there is communicative intent

> When discourse content is informative about goals/communicative intent, comprehenders use it to restrict focus alternatives.
theory 2: comprehension involves understanding communicative goals/speaker intent

> Ad hoc categories, categories based on world knowledge — rule out continuations that are unlikely given a real world situation (Barsalou [1982,1983])

> Expect different category effects depending on whether there is a salient goal, whether there is communicative intent

> When discourse content is informative about goals/communicative intent, comprehenders use it to restrict focus alternatives.

- ↑ category specificity = ↑ information = more restrictive
**predictions/implications**

1. **theory**: comprehension involves understanding communicative goals/speaker intent

   > *Ad hoc* categories, categories based on world knowledge — rule out continuations that are unlikely given a real world situation (Barsalou [1982,1983])

   > Expect different category effects depending on whether there is a salient goal, whether there is communicative intent

   > When discourse content is *informative* about goals/communicative intent, comprehenders use it to restrict focus alternatives.

   - $\uparrow$ category specificity $\Rightarrow$ $\uparrow$ information $\Rightarrow$ more restrictive
   - an informative sentence/context rules out more possibilities
You’re in a mall.
You’re in a shoe store.
Focus alternatives: conceptual category

**context sentence 1:**
1a. Lauren and Nathan are at the newsstand. ← [+Biasing context]
1b. Lauren and Nathan are at the drugstore. ← [−Biasing context]

**context sentence 2:**
2a. Lauren got some cigarettes and some magazines. ← [+Mention]
2b. Lauren got some cigarettes and some gum. ← [+Novel]

**target sentence:**
3a. Nathan got some magazines. ← [−Only]
3b. Nathan only got some magazines. ← [+Only]
Biasing context effect

![Graph showing the biasing context effect. The graph plots the ratio of target to target+competitor over time relative to the target word onset. Two lines represent wide context and narrow context, with error bars indicating variability.](image-url)
Timecourse of mention and bias effects
implications

theory2: comprehension involves understanding communicative goals/speaker intent

> When discourse content is *informative* about goals/communicative intent, comprehenders use it to restrict focus alternatives.
implications

theory2: comprehension involves understanding communicative goals/speaker intent

> When discourse content is *informative* about goals/communicative intent, comprehenders use it to restrict interpretation.
1 Narrow and broad explanations of the same data.

Case study 1. Category effects in the processing of alternative-sensitive particles

2 Making fast generalizations accurately.

Case study 2. Imprecision and cues to goal structure (a.k.a. Have you ever played flip-cup?)
roadmap

1 Narrow and broad explanations of the same data.

Case study 1. Category effects in the processing of alternative-sensitive particles

2 Making fast generalizations accurately.

Case study 2. Imprecision and cues to goal structure (a.k.a. Have you ever played flip-cup?)
New empirical domains

How do we generalize at the right level of abstraction?

overgeneralization ↔ overfitting

> Often data is sparse.
> Data is noisy, subject to selection bias, other non-interesting variability.
> We want to make broad generalizations (more universal = better).
> Linguistic data is really intricate/complex, so we gotta be careful.
New empirical domains

How do we generalize at the right level of abstraction?

overgeneralization $\leftrightarrow$ overfitting

> Data is REALLY sparse, or nonexistent.
> Data not only noisy; also preliminary, possibly mis-characterized.
> We want to make broad generalizations (more universal = better).
> Linguistic data is really intricate/complex, so we gotta be careful.
New empirical domains

How do we generalize at the right level of abstraction?

overgeneralization $\leftrightarrow$ overfitting

> Data is REALLY sparse, or nonexistent.
> Data not only noisy; also preliminary, possibly mis-characterized.
> We still want to make broad generalizations.
> Linguistic data is really intricate/complex, so we gotta be careful.
New empirical domains

How do we generalize at the right level of abstraction?

overgeneralization $\leftrightarrow$ overfitting

> Data is REALLY sparse, or nonexistent.
> Data not only noisy; also preliminary, possibly mis-characterized.
> We still want to make broad generalizations.
> Additional source of uncertainty: do we have the categories themselves right?
case study 2: gradable adjectives

**Hypothesis:** 2 subclasses of gradable adjectives differ in terms of how their standards are fixed in context

- relative gradable adjectives, standards of comparison
- absolute gradable adjectives, standards of precision

**Sources of uncertainty:**

- prior linguistic experience — items, subjects, items given subjects
- prior world experience — abstract representations of lexical/conceptual categories, implicit/indirect cues to speaker intent and/or goal structure
- are these the right categories?
uncertainty about the number and characterization of classes

2 classes?
uncertainty about the number and characterization of classes

1 class?  2 classes?
uncertainty about the number and characterization of classes

1 class? 2 classes? 3 classes?
relative gradable adjectives

Sameer is tall.

**Context-invariant part:**
Sameer’s height exceeds the standard of height determined by $\text{Comp}(C)$.

**Context-dependent part:**
$\text{Comp}(C) = \text{the salient comparison class}$
relative gradable adjectives

Sameer is tall.

**Context-invariant part:**
Sameer’s height exceeds the standard of height determined by $\text{Comp}(C)$.

**Context-dependent part:**
$\text{Comp}(C) = \text{the salient comparison class}$
relative gradable adjectives

My 6 yr old cousin is tall.

Context-invariant part:
Grace’s height in 1994 exceeds the standard of height determined by $\text{Comp}(C)$.

Context-dependent part:
$\text{Comp}(C) =$ the salient comparison class
The Sears Tower is **tall**.

**Context-invariant part:**
The Sears Tower’s height exceeds the standard of height determined by $\text{Comp}(C)$.

**Context-dependent part:**
$\text{Comp}(C) = \text{the salient comparison class}$
absolute gradable adjectives

My glass is **empty**.

**Context-invariant part:**
My glass’s emptiness exceeds the standard of emptiness determined by $\text{Comp}(C)$.

**Context-dependent part:**
$\text{Comp}(C) =$ the salient comparison class
absolute gradable adjectives

My glass is empty.

**Context-invariant part:**
My glass’s emptiness exceeds the standard of emptiness determined by $\text{Comp}(C)$.

**Context-dependent part:**
$\text{Comp}(C) = \text{the salient comparison class}$

[Diagram showing standard of comparison with a peak at 100%]
absolute gradable adjectives

The tank is **empty**.

**Context-invariant part:**
The tank’s emptiness exceeds the standard of emptiness determined by \( \text{Comp}(C) \).

**Context-dependent part:**
\( \text{Comp}(C) = \) the salient comparison class
The theater is **empty**.

**Context-invariant part:**
The theater’s emptiness exceeds the standard of emptiness determined by $\text{Comp}(C)$.

**Context-dependent part:**
$\text{Comp}(C) = \text{the salient comparison class}$
From: Centers of Gravity: In Conversation With an Ex-Gymnast:
I always felt that the short girls on the team were somehow ahead of the game.... But it was also luckily at a time where Svetlana Khorkina was getting a lot of attention.... She’s really tall.

Every time I make coffee in the afternoon, the pot doesn’t even finish brewing before (in one nanosecond) all my coworkers take a cup until it is empty! By the time I get to it, there is barely one cup left! They all wait until I brew it then proceed to take it all! Rarely is there a full cup of coffee left for me in the pot.
What is the nature of the indeterminacy w/absolute GAs?

**situation 1**
- high precision
  - completely empty

for use in a physics demo

**situation 2**
- low precision
  - emptyish

time to make more coffee

Chris Kim

Broad generalizations about pragmatic processing
What is the nature of the indeterminacy w/absolute GAs?

situation 1

- high precision
  - completely empty

for use in a physics demo

situation 2

- low precision
  - emptyish

- time to make more coffee

pragmatic context dependence

Strategy: Relative v. Absolute GAs

What we want to show:

- **Relative/Absolute** is the right distinction to make

Strategy:

- **Relative GAs**: Standard of comparison shows sensitivity to local discourse context

- **Absolute GAs**: Standard of precision sensitive to properties of global communicative context

- **Absolute but not Relative GAs** have ‘default’ precise meanings that emerge in the absence of strong contextual bias
Strategy: Relative v. Absolute GAs

What we want to show:

- Relative/Absolute is the right distinction to make

Strategy:

- **Relative GAs:** Standard of comparison shows sensitivity to local discourse context
- **Absolute GAs:** Standard of precision sensitive to properties of global communicative context
- **Absolute but not Relative GAs** have ‘default’ precise meanings that emerge in the absence of strong contextual bias
Homogeneity of world experience reduces one huge source of variability.

- more consensus about the superordinate category given good exemplars
- exemplars from different categories (or random pairings) → more categories assigned → inferences about communicative intent/goal structure based on highly variable cues (between subjects)
What category are these items examples of?

peas, carrots

Actual response: vegetables (× 16 votes)
What category are these items examples of?

peas, carrots
lemons, paper

Actual response:

- vegetables (16 votes)
- can be green
- can be yellow
- dryer sheets
- fish
- have pulp
- hurt if you get them in your eye
- lemonade stand
- nouns
- objects
- opposites
- secret writing
- sharp
- supplies
- tastes
- yellow (1 vote each)
### high v. low standard contexts

<table>
<thead>
<tr>
<th>relative GAs</th>
<th>absolute GAs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>high</strong></td>
<td>getting a kite that’s stuck on a chimney</td>
</tr>
<tr>
<td><strong>low</strong></td>
<td>getting a book from a high bookshelf</td>
</tr>
<tr>
<td>irrelevant</td>
<td></td>
</tr>
</tbody>
</table>
### high v. low standard contexts

<table>
<thead>
<tr>
<th>relative GAs</th>
<th>absolute GAs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>high</strong></td>
<td><strong>organizing a</strong></td>
</tr>
<tr>
<td>getting a kite that’s stuck on a chimney</td>
<td>game of flip-cup</td>
</tr>
<tr>
<td><strong>low</strong></td>
<td><strong>refilling beverages at a party</strong></td>
</tr>
<tr>
<td>getting a book from a high bookshelf</td>
<td></td>
</tr>
<tr>
<td>irrelevant</td>
<td></td>
</tr>
</tbody>
</table>
### high v. low standard contexts

<table>
<thead>
<tr>
<th>relative GAs</th>
<th>absolute GAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>high</td>
</tr>
<tr>
<td>getting a kite that’s stuck on a chimney</td>
<td>organizing a game of flip-cup</td>
</tr>
<tr>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>getting a book from a high bookshelf</td>
<td>refilling beverages at a party</td>
</tr>
<tr>
<td>irrelevant</td>
<td>irrelevant</td>
</tr>
<tr>
<td>looking for a bike</td>
<td>looking for a bike</td>
</tr>
</tbody>
</table>
### high v. low standard contexts

<table>
<thead>
<tr>
<th></th>
<th>relative GAs</th>
<th>absolute GAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>getting a kite that’s stuck on a chimney</td>
<td>organizing a game of flip-cup</td>
</tr>
<tr>
<td>low</td>
<td>getting a book from a high bookshelf</td>
<td>refilling beverages at a party</td>
</tr>
<tr>
<td>irrelevant</td>
<td>looking for a bike</td>
<td>looking for a bike</td>
</tr>
</tbody>
</table>

- tall ladder?  
- empty cup?
High standard context

Duncan is trying to get his kite, which is caught on the chimney. He’s looking for a tall ladder.
High standard context

Duncan is trying to get his kite, which is caught on the chimney. He’s looking for a tall ladder.

This ladder is tall.

☐ yes
☐ no
extremely tall/completely empty

High standard context

Duncan is trying to get his kite, which is caught on the chimney. He’s looking for a tall ladder.

This ladder is tall.

☐ yes
☐ no
Duncan is trying to reach a book on the top shelf of the bookcase. He's looking for a tall ladder.
Low standard context

Duncan is trying to reach a book on the top shelf of the bookcase. He’s looking for a tall ladder.

This ladder is tall.

☑️ yes
☐ no
High precision context

Max is organizing a game of flip-cup at the party.
He’s looking for empty cups.
High precision context

Max is organizing a game of flip-cup at the party.
He’s looking for empty cups.

This cup it empty.

☑ yes
☐ no
Low precision context

Max is going to the kitchen to refill his beverage at the party. He’s asking if anyone else has an empty cup.
extremely tall/completely empty

Low precision context

Max is going to the kitchen to refill his beverage at the party. He’s asking if anyone else has an empty cup.

This cup it empty.

☐ yes
☐ no
Duncan is trying to get his kite, which is caught on the chimney. He’s looking for a tall ladder.

This ladder is tall.

☐ yes?
☐ no?
Duncan is trying to reach a book on the top shelf of the bookcase. He’s looking for a tall ladder.

This ladder is tall.

☐ yes?
☐ no?
High precision context

Max is organizing a game of flip-cup at the party.
He’s looking for empty cups.

This cup it empty.

☐ yes?
☐ no?
Max is going to the kitchen to refill his beverage at the party. He’s asking if anyone else has an empty cup.
Irrelevant context

Duncan missed his bus in the morning, so he decided to bike to school and went looking for his bike. It was in the garage next to the ladder.

This ladder is tall.

☑  yes
☐  no
Duncan missed his bus in the morning, so he decided to bike to school and went looking for his bike. It was in the garage next to the ladder.

This ladder is tall.

☐ yes ?
☐ no ?
irrelevant contexts — extremely tall/completely empty

irrelevant context

After Max got home from the picnic, he realized he didn’t have his cell phone.
It turns out he left it on one of the picnic tables, where it was hidden from view by some cups.

this cup is empty.

☐ yes

☐ no
irrelevant contexts — *pretty tall/pretty empty*

**irrelevant context**

After Max got home from the picnic, he realized he didn’t have his cell phone.
It turns out he left it on one of the picnic tables, where it was hidden from view by some cups.

This cup is empty.

☐ yes ?
☐ no ?
Compare accept rates for: (i) high standard context, (ii) low standard context, (iii) irrelevant context

extreme/perfect exemplar: accept rates at ceiling — no differences between conditions
extremely tall/completely empty
Compare accept rates for: (i) high standard context, (ii) low standard context, (iii) irrelevant context

extreme/perfect exemplar: accept rates at ceiling — no differences between conditions

non-extreme exemplar:
  ■ relative GAs: general increase in accept rate from high standard to low standard to irrelevant
  ■ absolute GAs:
    ■ increase in accept rate from high standard to low standard
    ■ irrelevant context matches high standard accept rate
Broad generalizations about pragmatic processing
goal sensitivity and ‘pragmatic context dependence’?

**Lexicon**

form1 ↔ meaning1
form2 ↔ meaning2
form3 ↔ meaning3
...

**In context**

form3 ←?→ meaning3a
←?→ meaning3b
←?→ meaning3c

The tank is *empty*.
goal sensitivity and ‘pragmatic context dependence’?

Lexicon

form1 $\leftrightarrow$ meaning1
form2 $\leftrightarrow$ meaning2
form3 $\leftrightarrow$ meaning3
...

In context

empty $\leftarrow$ ? $\rightarrow$ 100% empty
$\leftarrow$ ? $\rightarrow$ might as well be empty
$\leftarrow$ ? $\rightarrow$ ...

The tank is empty.
roadmap

1. Narrow and broad explanations of the same data.

Case study 1. Category effects in the processing of alternative-sensitive particles


Case study 2. Imprecision and cues to goal structure (a.k.a. Have you ever played flip-cup?)
How do we generalize at the right level of abstraction?

\[ \text{overgeneralization} \leftrightarrow \text{overfitting} \]

Broad generalizations:

- framework for cue combination
- tradeoffs between accuracy/parsimony
- communicative function (what’s different about processing information for communication and general information processing?)

Predictions about dimensions of cross-cultural variation.
Thanks!

Especially taxpayers!:  
NSF BCS-1227144