Grammatical Knowledge is Fundamentally Probabilistic

Roger Levy

UC San Diego
Department of Linguistics
Annual Meeting of the Linguistic Society of America

10 January 2015
A couple of central questions in linguistics:
A couple of central questions in linguistics:

- How do we know when relative preferences among strings genuinely reflect grammatical knowledge ("competence")?
A couple of central questions in linguistics:

- How do we know when relative preferences among strings genuinely reflect grammatical knowledge ("competence")?
- Is the content of grammatical knowledge purely categorical, or does it include a probabilistic component?
Introduction

¬ Desiderata for explaining relative preferences through appeal to grammatical knowledge
Introduction

- Desiderata for explaining relative preferences through appeal to grammatical knowledge
  - The preferences cannot be accounted for through extralinguistic knowledge or independently motivated performance constraints
Introduction

- Desiderata for explaining relative preferences through appeal to grammatical knowledge
  - The preferences cannot be accounted for through extralinguistic knowledge or independently motivated performance constraints
  - Hypothesizing the grammatical knowledge improves descriptive adequacy
Desiderata for explaining relative preferences through appeal to grammatical knowledge

- The preferences cannot be accounted for through extralinguistic knowledge or independently motivated performance constraints
- Hypothesizing the grammatical knowledge improves descriptive adequacy

What types of relative preferences might reflect probabilistic, or gradient, knowledge?
Introduction

- Desiderata for explaining relative preferences through appeal to grammatical knowledge
  - The preferences cannot be accounted for through extralinguistic knowledge or independently motivated performance constraints
  - Hypothesizing the grammatical knowledge improves descriptive adequacy
- What types of relative preferences might reflect probabilistic, or gradient, knowledge?
  - If, among structures that are grammatically licensed, some are “better” than (reliably preferred to) others
Today’s argument

Does the content of grammatical knowledge include a probabilistic component?

- In the domain of syntax, this question remains unresolved
Today’s argument

Does the content of grammatical knowledge include a probabilistic component?

- In the domain of syntax, this question remains unresolved
- (In phonology, a clear “yes” —Hayes, Albright, Wilson, . . . )
Today’s argument

Does the content of grammatical knowledge include a probabilistic component?

- In the domain of syntax, this question remains unresolved
- (In phonology, a clear “yes” —Hayes, Albright, Wilson, . . .)
- Spirited “yes” arguments have been made, but they’re not conclusive
Today’s argument

Does the content of grammatical knowledge include a probabilistic component?

- In the domain of syntax, this question remains unresolved
- (In phonology, a clear “yes” —Hayes, Albright, Wilson, . . .)
- Spirited “yes” arguments have been made, but they’re not conclusive
- I will adduce a (hopefully) stronger argument
**Argument from acquisition**

**Argument:** learning of probabilistic grammars (Horning, 1969) may be more tractable than learning of non-probabilistic grammars (Gold, 1967) (Manning, 2003; Perfors et al., 2011; Clark and Lappin, 2009)
Argument from acquisition

**Argument:** learning of probabilistic grammars (Horning, 1969) may be more tractable than learning of non-probabilistic grammars (Gold, 1967) (Manning, 2003; Perfors et al., 2011; Clark and Lappin, 2009)

**Objections:**

- Maybe learning bias (a.k.a. UG) leaves the learning problem minimally difficult
Argument from acquisition

**Argument:** learning of probabilistic grammars (Horning, 1969) may be more tractable than learning of non-probabilistic grammars (Gold, 1967) (Manning, 2003; Perfors et al., 2011; Clark and Lappin, 2009)

**Objections:**

- Maybe learning bias (a.k.a. UG) leaves the learning problem minimally difficult
- Also, non-probabilistic grammar induction is getting better and better (e.g., Clark and Eyraud, 2006)
Argument from acquisition

**Argument:** learning of probabilistic grammars (Horning, 1969) may be more tractable than learning of non-probabilistic grammars (Gold, 1967) (Manning, 2003; Perfors et al., 2011; Clark and Lappin, 2009)

**Objections:**

- Maybe learning bias (a.k.a. UG) leaves the learning problem minimally difficult
- Also, non-probabilistic grammar induction is getting better and better (e.g., Clark and Eyraud, 2006)
- Or, probabilities may be used during acquisition and then discarded in encoding a categorical competence grammar
**Argument:** there are systematic attachment preferences for configurations such as

\[
[\text{VP } V [\text{VP } V \ldots ] \ldots ] \text{ Adv}
\]

e.g.,

*John said that Mary left yesterday*

and those preferences match corpus frequencies (Jurafsky, 1996, 2003; Levy, 2008; MacDonald and Thornton, 2009)
Argument: there are systematic attachment preferences for configurations such as

\[ [\text{VP} \ V \ [\text{VP} \ V \ldots] \ldots] \text{ Adv} \]

e.g.,

*John said that Mary left yesterday*

and those preferences match corpus frequencies (Jurafsky, 1996, 2003; Levy, 2008; MacDonald and Thornton, 2009)

Objection: This type of attachment preference may be due to linear-distance locality preferences (Hawkins, 1994; Gibson, 1998), not structural frequencies
**Argument:** For a given set of thematic roles, different verbs seem to have idiosyncratically different preferences for syntactic realization (Manning, 2003; Bresnan et al., 2007):

The dative alternation

\[
\begin{align*}
&[V \ NP_{\text{Goal}} \ NP_{\text{Theme}}]\text{-preferring} \\
&\quad \text{tell} \\
&\quad \text{teach} \\
&\quad \text{charge} \\
&\quad \text{show} \\
&[V \ NP_{\text{Theme}} \ PP_{\text{Goal}}]\text{-preferring} \\
&\quad \text{loan} \\
&\quad \text{bring} \\
&\quad \text{sell} \\
&\quad \text{take}
\end{align*}
\]
Lexeme-specific syntactic preference

**Argument:** For a given set of thematic roles, different verbs seem to have idiosyncratically different preferences for syntactic realization (Manning, 2003; Bresnan et al., 2007):

**The dative alternation**

<table>
<thead>
<tr>
<th>[V NP$<em>{\langle \text{Goal} \rangle}$ NP$</em>{\langle \text{Theme} \rangle}$]-preferring</th>
<th>[V NP$<em>{\langle \text{Theme} \rangle}$ PP$</em>{\langle \text{Goal} \rangle}$]-preferring</th>
</tr>
</thead>
<tbody>
<tr>
<td>tell</td>
<td>loan</td>
</tr>
<tr>
<td>teach</td>
<td>bring</td>
</tr>
<tr>
<td>charge</td>
<td>sell</td>
</tr>
<tr>
<td>show</td>
<td>take</td>
</tr>
</tbody>
</table>

**Objection:** Lexeme-specific preferences may be epiphenomena of differences in *shades of meaning* among syntactic realizations
The internal structure of coordinate categories

- Principle of *Conjoin Likes* (Chomsky, 1965)
  \[ X \rightarrow X \text{ Conj } X \]
The internal structure of coordinate categories

- Principle of *Conjoin Likes* (Chomsky, 1965)
  \[ X \rightarrow X \text{ Conj } X \]

- Empirically false for gross syntactic category (Sag et al., 1985):
  
  *Pat is a Republican and proud of it (NP and AdjP)*
The internal structure of coordinate categories

- Principle of *Conjoin Likes* (Chomsky, 1965)
  \[ X \rightarrow X \text{ Conj } X \]

- Empirically false for gross syntactic category (Sag et al., 1985):
  
  *Pat is a Republican and proud of it (NP and AdjP)*
The internal structure of coordinate categories

- Principle of *Conjoin Likes* (Chomsky, 1965)
  \[ X \rightarrow X \text{ Conj } X \]

- Empirically false for gross syntactic category (Sag et al., 1985):
  
  *Pat is a Republican and proud of it* (NP and AdjP)
  *That was a rude remark and in very bad taste* (NP and PP)
The internal structure of coordinate categories

- Principle of *Conjoin Likes* (Chomsky, 1965)
  \[ X \rightarrow X \text{ Conj } X \]

- Empirically false for gross syntactic category (Sag et al., 1985):
  
  *Pat is a Republican and proud of it (NP and AdjP)*
  *That was a rude remark and in very bad taste (NP and PP)*

- So... was “Conjoin Likes” just wrong?
The internal structure of coordinate categories

- Principle of *Conjoin Likes* (Chomsky, 1965)
  \[ X \rightarrow X \text{ Conj } X \]

- Empirically false for gross syntactic category (Sag et al., 1985):
  
  *Pat is a Republican and proud of it* (NP and AdjP)
  *That was a rude remark and in very bad taste* (NP and PP)

- So... was “Conjoin Likes” just wrong?

- Is there anything left for grammar to say about a “tendency” for coordinated categories to be like one another?
Unlike-category coordinations are easy to find in corpora.

His son had been friendly, a big fellow of fifty or more, a fishing-boat captain and powerful like the sea

(Parsed Brown corpus)
Unlike-category coordinations are easy to find in corpora:

His son had been friendly, a big fellow of fifty or more, a fishing-boat captain and powerful like the sea.

But there is a huge quantitative tendency for coordination to be of like categories in corpora:

<table>
<thead>
<tr>
<th>Right-hand conjunct</th>
<th>1308</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left-hand conjunct</td>
<td>6</td>
<td>114</td>
</tr>
</tbody>
</table>

(Parsed Brown corpus)
Conjoin Likes as a gradient grammatical constraint?

- Perhaps this is gradient Conjoin Likes!
**Conjoin Likes** as a gradient grammatical constraint?

- Perhaps this is gradient *Conjoin Likes*!
- But should we really attribute this to the grammar proper?
Conjoin Likes as a gradient grammatical constraint?

- Perhaps this is gradient Conjoin Likes!
- But should we really attribute this to the grammar proper?
- We need:
Conjoin Likes as a gradient grammatical constraint?

- Perhaps this is gradient Conjoin Likes!
- But should we really attribute this to the grammar proper?
- We need:
  - A precise proposal of gradient Conjoin Likes
Conjoin Likes as a gradient grammatical constraint?

- Perhaps this is gradient Conjoin Likes!
- But should we really attribute this to the grammar proper?
- We need:
  - A precise proposal of gradient Conjoin Likes
  - Evidence of relative preference for like-conjunct coordinations that controls for meaning and external context
Conjoin Likes as a gradient grammatical constraint?

A cautionary thought experiment:

```
S
  NP  VP
    V_{trans}  X
        Y and Z

S
  NP  VP
    be  X
        Y and Z

NP
  X  N
    Y and Z
```

Tendency for Conjoin Likes
Conjoin Likes as a gradient grammatical constraint?

A cautionary thought experiment:

$$\begin{array}{c}
\text{S} \\
\text{NP} \quad \text{VP} \\
\quad \text{\(V_{\text{trans}}\)} \quad \text{X} \\
\quad \text{Y and Z} \\
\downarrow \\
\text{NP and NP} \\
\end{array}$$

$$\begin{array}{c}
\text{S} \\
\text{NP} \quad \text{VP} \\
\quad \text{be} \quad \text{X} \\
\quad \text{Y and Z} \\
\downarrow \\
\text{Uncorrelated Mixture} \\
\end{array}$$

$$\begin{array}{c}
\text{NP} \\
\text{X} \quad \text{N} \\
\downarrow \\
\text{AdjP and AdjP} \\
\end{array}$$

$$\text{NP and NP}$$
Conjoin Likes as a gradient grammatical constraint?

A cautionary thought experiment:

```
S
  NP   VP
    Vtrans X
      Y and Z

S
  NP   VP
    be X
      Y and Z

NP
  X N
    Y and Z
```

NP and NP
Uncorrelated Mixture
AdjP and AdjP

Tendency for Conjoin Likes
What “gradient coordination of like categories” means:

\[ P(F_1, F_2|M_1, M_2) \]

is especially high when \( F_1 \) and \( F_2 \) are “like” in the traditional sense of

\[ X \rightarrow X \text{ and } X \]

Fully technically:

\[
\text{pMI}(F_1, F_2|M_1, M_2) = \log \frac{P(F_1, F_2|M_1, M_2)}{P(F_1|M_1)P(F_2|M_2)}
\]

is monotonically increasing in the structural similarity of \( F_1 \) and \( F_2 \)
Empirical prediction

If more probable forms are gradiently “more grammatical” to the native speaker...
Empirical prediction

If more probable forms are gradiently “more grammatical” to the native speaker... then like-category coordinations should be judged more natural, or acceptable, than unlike-category coordinations.
Empirical prediction

If more probable forms are gradiently “more grammatical” to the native speaker... then like-category coordinations should be judged more natural, or acceptable, than unlike-category coordinations.

More precisely, if

\[
\left\{ \begin{array}{l}
pMI(\text{AdjP}, \text{AdjP}|M_1, M_2) \\
pMI(\text{NP}, \text{NP}|M_1, M_2)
\end{array} \right\} > \left\{ \begin{array}{l}
pMI(\text{AdjP}, \text{NP}|M_1, M_2) \\
pMI(\text{NP}, \text{AdjP}|M_1, M_2)
\end{array} \right\}
\]

then

\[
\log P(\text{AdjP}, \text{AdjP}|M_1, M_2) + \log P(\text{NP}, \text{NP}|M_1, M_2) \geq \log P(\text{AdjP}, \text{NP}|M_1, M_2) + \log P(\text{NP}, \text{AdjP}|M_1, M_2)
\]

and like-conjunct coordinations should be superadditively good.
Experiment 1

Acceptability judgment study (scale of 1–9):

*Pat is a Republican and a freak.* [Noun Noun]

*Pat is a Republican and freaky.* [Noun Adj]

*Pat is Republican and a freak.* [Adj Noun]

*Pat is Republican and freaky.* [Adj Adj]

(Baseline: *The children decorated the sparkling ornaments onto the tree was a 4.*)

(All experiments: Mechanical Turk native speakers, 12 items, the same 21 fillers (all “good”), ≥ 75 subjects.)
Experiment 1: Results

The gradient preference for coordination of unlike categories is pretty strong!
Greater explanatory power of gradient constraints

- We saw that “Conjoin Likes” is categorically false, but “probabilistically” true
Greater explanatory power of gradient constraints

- We saw that “Conjoin Likes” is categorically false, but “probabilistically” true
- But why stop at major syntactic categories—what about category-*internal* structure (Johnson, 1998; Klein and Manning, 2003)?
Greater explanatory power of gradient constraints

- We saw that “Conjoin Likes” is categorically false, but “probabilistically” true.
- But why stop at major syntactic categories—what about category-internal structure (Johnson, 1998; Klein and Manning, 2003)?
- Such a grammatical preference has previously been explored under the rubric of PARALLELISM (Frazier et al., 1984; Hale et al., 2006; Dubey et al., 2008)
NP-internal parallelism: the genitive alternation

<table>
<thead>
<tr>
<th>Postnominal</th>
<th>Prenominal</th>
</tr>
</thead>
<tbody>
<tr>
<td>The future of our country</td>
<td>Our country’s future</td>
</tr>
<tr>
<td>The base of the lamp</td>
<td>The lamp’s base</td>
</tr>
<tr>
<td>The tail of a cat</td>
<td>A cat’s tail</td>
</tr>
</tbody>
</table>
NP-internal parallelism: the genitive alternation

Postnominal
The future of our country
The base of the lamp
The tail of a cat

Prenominal
Our country’s future
The lamp’s base
A cat’s tail

\[
\begin{array}{c}
\text{NP} \\
\text{NP} \quad \text{Conj} \quad \text{NP} \\
\text{NP} \quad \text{PP}_{\text{of}} \quad \text{and} \quad \text{NP} \quad \text{PP}_{\text{of}} \\
\text{NP} \quad \text{PP}_{\text{of}} \quad \text{and} \quad \text{NP} \quad \text{PP}_{\text{of}} \\
\end{array}
\]
NP-internal parallelism: the genitive alternation

Postnominal
The future of our country
The base of the lamp
The tail of a cat

Prenominal
~ Our country’s future
~ The lamp’s base
~ A cat’s tail

Diagram:

NP
  ┌─ NP
  │   └── Conj
  │       ┌─ NP
  │         │   └── PP_{of}
  │       └─ and └─ NP
  │                   └── PP_{of}
  └── NP

NP
  ┌─ NP
  │   └── PossP
  │       └── NP
  │       ┌─ Conj
  │         │   └─ and └─ PP_{of}
  │         └── NP
  └── NP

NP
  ┌─ NP
  │   └── PossP
  │       └── NP
  │       ┌─ Conj
  │         │   └─ and └─ PossP
  │         └── NP
  └── NP
Corpus data on genitive alternation parallelism

<table>
<thead>
<tr>
<th>Left Conjunct</th>
<th>Right Conjunct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Post</td>
</tr>
<tr>
<td>Post</td>
<td>77</td>
</tr>
<tr>
<td>Pre</td>
<td>20</td>
</tr>
</tbody>
</table>

- There is also strong evidence for a parallelism preference in the genitive alternation...
There is also strong evidence for a parallelism preference in the genitive alternation. ... But once again this analysis fails to control for external context and conjunct meanings $M_1, M_2$.
Corpus data on genitive alternation parallelism

<table>
<thead>
<tr>
<th>Left Conjunct</th>
<th>Right Conjunct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Post</td>
</tr>
<tr>
<td>Post</td>
<td>77</td>
</tr>
<tr>
<td>Pre</td>
<td>20</td>
</tr>
</tbody>
</table>

- There is also strong evidence for a parallelism preference in the genitive alternation...
- ...but once again this analysis fails to control for external context and conjunct meanings $M_1, M_2$
- We can control this more tightly with an experiment
Experiment 2: genitive parallelism

Acceptability judgment study (scale of 1–9):

Terry assembled...

...the frame of the chair and the base of the lamp. [Post Post]

...the frame of the chair and the lamp’s base. [Post Pre ]

...the chair’s frame and the base of the lamp. [Pre Post]

...the chair’s frame and the lamp’s base. [Pre Pre ]

(Baseline: The children decorated the sparkling ornaments onto the tree was a 4.)
There is also a preference for parallelism among realizations of the genitive alternation!
Searching for deeper parallelism effects
Searching for deeper parallelism effects

Like categories!

\[
\begin{align*}
\text{XP} & \\
?P & \text{Conj} & ?P \\
& \text{and}
\end{align*}
\]
Searching for deeper parallelism effects

Like categories!

XP

?P Conj ?P

and

Genitive parallelism!

NP

NP Conj NP

Searching for deeper parallelism effects

Like categories!

Genitive parallelism!

Embedded-genitive parallelism???
Experiment 3: parallelism of embedded genitive

Kim is full of...

...skepticism about the future of our country and pessimism regarding the prospects of our children.
...skepticism about the future of our country and pessimism regarding our children’s prospects.
...skepticism about our country’s future and pessimism regarding the prospects of our children.
...skepticism about our country’s future and pessimism regarding our children’s prospects.
This deep in a coordinate NP structure, the parallelism preference has disappeared!
Ruling out priming

Argument thus far:

§ Categorical
§ Conjoin Likes does not exist
§ But gradient Conjoin Likes does
§ We've ruled out external context and conjunct meaning as confounds
§ Conjoin Likes constitutes non-categorical, probabilistic grammatical knowledge

Possible objection: the gradient Conjoin Likes preference has another, independently motivated explanation

Proposed gradient knowledge
Independent explanation
Local attachment preference (said. . . left. . . tomorrow)
Memory-driven locality
Verb-specific argument realization preferences
Constructional meaning differences
Gradient Conjoin Likes Syntactic priming?
Ruling out priming

Argument thus far:

- Categorical Conjoin Likes does not exist
Ruling out priming

Argument thus far:
  - Categorical *Conjoin Likes* does not exist
  - But gradient *Conjoin Likes* does
Ruling out priming

Argument thus far:

- Categorical *Conjoin Likes* does not exist
- But gradient *Conjoin Likes* does
- We’ve ruled out external context and conjunct meaning as confounds
Ruling out priming

Argument thus far:

- Categorical *Conjoin Likes* does not exist
- But gradient *Conjoin Likes* does
- We’ve ruled out external context and conjunct meaning as confounds
- *Conjoin Likes* constitutes non-categorical, probabilistic grammatical knowledge
Ruling out priming

Argument thus far:

- Categorical *Conjoin Likes* does not exist
- But gradient *Conjoin Likes* does
- We’ve ruled out external context and conjunct meaning as confounds
- *Conjoin Likes* constitutes non-categorical, probabilistic grammatical knowledge

Possible objection: the gradient *Conjoin Likes* preference has another, independently motivated explanation

Proposed gradient knowledge

- Independent explanation: Local attachment preference
  - (said. . . left. . . tomorrow)
- Memory-driven locality
- Verb-specific argument realization
- Constructional meaning differences
- *Conjoin Likes* Syntactic priming?
Ruling out priming

Argument thus far:

- Categorical *Conjoin Likes* does not exist
- But gradient *Conjoin Likes* does
- We’ve ruled out external context and conjunct meaning as confounds
- *Conjoin Likes* constitutes non-categorical, probabilistic grammatical knowledge

**Possible objection:** the gradient *Conjoin Likes* preference has another, independently motivated explanation
Ruling out priming

Argument thus far:

- Categorical *Conjoin Likes* does not exist
- But gradient *Conjoin Likes* does
- We’ve ruled out external context and conjunct meaning as confounds
- *Conjoin Likes* constitutes non-categorical, probabilistic grammatical knowledge

Possible objection: the gradient *Conjoin Likes* preference has another, independently motivated explanation

Proposed gradient knowledge Independent explanation
Ruling out priming

Argument thus far:

- Categorical *Conjoin Likes* does not exist
- But gradient *Conjoin Likes* does
- We’ve ruled out external context and conjunct meaning as confounds
- *Conjoin Likes* constitutes non-categorical, probabilistic grammatical knowledge

Possible objection: the gradient *Conjoin Likes* preference has another, independently motivated explanation

**Proposed gradient knowledge**    **Independent explanation**
Local attachment preference    Memory-driven locality
*(said... left... tomorrow)*
Ruling out priming

Argument thus far:

- Categorical Conjoin Likes does not exist
- But gradient Conjoin Likes does
- We’ve ruled out external context and conjunct meaning as confounds
- Conjoin Likes constitutes non-categorical, probabilistic grammatical knowledge

Possible objection: the gradient Conjoin Likes preference has another, independently motivated explanation

Proposed gradient knowledge
Local attachment preference
(said... left... tomorrow)
Verb-specific argument realization preferences
Gradient Conjoin Likes

Independent explanation
Memory-driven locality
Constructional meaning differences
Syntactic priming?
Subject/object configuration rather than coordinated object NPs:

The future of our country will determine the prospects of our children. [Post Post]

The future of our country will determine our children’s prospects. [Post Pre]

Our country’s future will determine the prospects of our children. [Pre Post]

Our country’s future will determine our children’s prospects. [Pre Pre]
There is a parallelism effect here! So...?
Strength of parallelism across 4 experiments

- Unlike Categories
- Genitive Alternation
- Embedded Genitive
- Subject/Object Genitive

Parallelism effect size

* n.s.

Conjoin Likes isn't just priming!
**Strength of parallelism across 4 experiments**

- **Unlike Categories**
  - Parallelism effect size: 0.6
  - Note: *n.s.*

- **Genitive Alternation**
  - Parallelism effect size: 0.4

- **Embedded Genitive**
  - Parallelism effect size: 0.2

- **Subject/Object Genitive**
  - Parallelism effect size: 0.0

*Conjoin Likes isn’t just priming!*
Strength of parallelism across 4 experiments

- Unlike Categories: 0.8
- Genitive Alternation: 0.6
- Embedded Genitive: 0.2
- Subject/Object Genitive: 0.0

* Conjoin Likes isn't just priming!
Strength of parallelism across 4 experiments

Comparison of Parallelism effect size:

- Unlike Categories: 0.6 ± 0.1
- Genitive Alternation: 0.3 ± 0.05
- Embedded Genitive: n.s.
- Subject/Object Genitive: 0.1 ± 0.05

* indicates a significant effect.
Strength of parallelism across 4 experiments

Parallelism effect size

Like
Categories
Genitive
Alternation
Embedded
Genitive
Subject/Object
Genitive

Conjoin Likes isn’t just priming!

* n.s.
Grammar has very little to say about *categorical* constraints on the relation between conjuncts.
Grammar has very little to say about **categorical** constraints on the relation between conjuncts.

But corpus data suggest there’s much more to say about the **gradient** constraints on their relation.
Discussion

Grammar has very little to say about categorical constraints on the relation between conjuncts.

But corpus data suggest there’s much more to say about the gradient constraints on their relation.

We now have the technical tools to formally characterize these gradient constraints.
Grammar has very little to say about categorical constraints on the relation between conjuncts.

But corpus data suggest there’s much more to say about the gradient constraints on their relation.

We now have the technical tools to formally characterize these gradient constraints.

This formalization revealed a weakness of (sparse) corpus data and guided experiments to test for and quantify the strength of these constraints.
Grammar has very little to say about **categorical** constraints on the relation between conjuncts. But corpus data suggest there’s much more to say about the **gradient** constraints on their relation. We now have the technical tools to formally characterize these gradient constraints. This formalization revealed a weakness of (sparse) corpus data and guided experiments to test for and quantify the strength of these constraints. We found that gradient “Conjoin Likes” is real, and has greater explanatory reach than was ever claimed for the categorical version!
Conjoin Likes as probabilistic grammatical knowledge

- Content of gradient Conjoin Likes:
Conjoin Likes as probabilistic grammatical knowledge

- Content of gradient *Conjoin Likes*:
  - Pointwise mutual information (pMI) is higher the more structurally similar the form of the conjuncts
Conjoin Likes as probabilistic grammatical knowledge

- Content of gradient Conjoin Likes:
  - Pointwise mutual information (pMI) is higher the more structurally similar the form of the conjuncts
  - Sensitivity of pMI to conjunct similarity falls off as structural locality decreases
Conjoin Likes as probabilistic grammatical knowledge

- Content of gradient Conjoin Likes:
  - Pointwise mutual information (pMI) is higher the more structurally similar the form of the conjuncts
  - Sensitivity of pMI to conjunct similarity falls off as structural locality decreases

- These generalizations cannot be reduced to real-world knowledge or independently motivated performance constraints
**Conjoin Likes as probabilistic grammatical knowledge**

- Content of gradient *Conjoin Likes*:
  - Pointwise mutual information (pMI) is higher the more structurally similar the form of the conjuncts
  - Sensitivity of pMI to conjunct similarity falls off as structural locality decreases

- These generalizations cannot be reduced to real-world knowledge or independently motivated performance constraints

- Gradient *Conjoin Likes*—and perhaps other unlexicalized and lexicalized syntactic preferences—constitutes fundamentally probabilistic grammatical knowledge


References III


