Are formal restrictions on crossing dependencies epiphenomenal?

Himanshu Yadav  
Indian Institute of Technology, Kanpur  
yadavhimanshu059@gmail.com

Samar Husain*  
Indian Institute of Technology, Delhi  
samar@hss.iitd.ac.in

Richard Futrell*  
University of California, Irvine  
@rljfutrell  
rfutrell@uci.edu

18th International Workshop on Treebanks & Linguistic Theory  
2019-08-28
Dependency trees are usually projective
Dependency trees are usually projective

I know you think you saw Bob yesterday.
Dependency trees are usually projective

I know you think you saw Bob yesterday.

(that is, the lines don’t cross)
But sometimes they’re not

I know who you think you saw yesterday.
What do we know about crossing dependencies?
What do we know about crossing dependencies?

- They are rare (Ferrer-i-Cancho et al., 2018).
What do we know about crossing dependencies?

- They are rare (Ferrer-i-Cancho et al., 2018).
- They occur much less frequently than would be expected in random reorderings of real dependency trees.
What do we know about crossing dependencies?

• They are rare (Ferrer-i-Cancho et al., 2018).
  • They occur much less frequently than would be expected in random reorderings of real dependency trees.
• They appear to be formally restricted.
What do we know about crossing dependencies?

- They are **rare** (Ferrer-i-Cancho et al., 2018).
  - They occur much less frequently than would be expected in random reorderings of real dependency trees.
- They appear to be **formally restricted**.
  - E.g., ~96% of natural language structures in dependency treebanks are **well-nested with gap degree <2** (Kuhlmann, 2013).
What do we know about crossing dependencies?

- They are rare (Ferrer-i-Cancho et al., 2018).
  - They occur much less frequently than would be expected in random reorderings of real dependency trees.
- They appear to be formally restricted.
  - E.g., ~96% of natural language structures in dependency treebanks are well-nested with gap degree <2 (Kuhlmann, 2013).
  - ~96% of dependency trees are <= 1-end-point-crossing (Pitler et al., 2013)
What do we know about crossing dependencies?

- They are rare (Ferrer-i-Cancho et al., 2018).
  - They occur much less frequently than would be expected in random reorderings of real dependency trees.
- They appear to be formally restricted.
  - E.g., ~96% of natural language structures in dependency treebanks are well-nested with gap degree <2 (Kuhlmann, 2013).
  - ~96% of dependency trees are <= 1-end-point-crossing (Pitler et al., 2013)
  - And more…
What do we know about crossing dependencies?

- They are rare (Ferrer-i-Cancho et al., 2018).
  - They occur much less frequently than would be expected in random reorderings of real dependency trees.
- They appear to be formally restricted.
  - E.g., ~96% of natural language structures in dependency treebanks are well-nested with gap degree <2 (Kuhlmann, 2013).
  - ~96% of dependency trees are <= 1-end-point-crossing (Pitler et al., 2013)
  - And more…
- Many formal restrictions have been proposed in the literature. We call these formal constraints crossing constraints.
What do we know about crossing dependencies?

- They are **rare** (Ferrer-i-Cancho et al., 2018).
  - They occur much less frequently than would be expected in random reorderings of real dependency trees.
- They appear to be **formally restricted**.
  - E.g., ~96% of natural language structures in dependency treebanks are **well-nested with gap degree <2** (Kuhlmann, 2013).
  - ~96% of dependency trees are <= **1-end-point-crossing** (Pitler et al., 2013)
  - And more…
- Many formal restrictions have been proposed in the literature. We call these formal constraints **crossing constraints**.
Why do we care about crossing constraints?
Why do we care about crossing constraints?

• They are important for designing efficient parsers. For example:
Why do we care about crossing constraints?

- They are important for designing efficient parsers. For example:
  - Well-nested dependency trees with gap degree $k$ can be parsed in time $O(n^{2k+3})$ (Gómez-Rodríguez et al., 2011).
Why do we care about crossing constraints?

- They are important for designing efficient parsers. For example:
  - **Well-nested** dependency trees with gap degree $k$ can be parsed in time $O(n^{2k+3})$ (Gómez-Rodríguez et al., 2011).
  - Dependency trees that are **1-end-point-crossing** (a subset of **2-planar** trees) can be parsed in time $O(n^4)$ (Pitler et al., 2013).
Why *do we care* about crossing constraints?
Why do we care about crossing constraints?

- They **delimit the formal language class** of human language (Chomsky & Schutzenberger, 1963).
Why do we care about crossing constraints?

- They **delimit the formal language class** of human language (Chomsky & Schutzenberger, 1963).
Why do we care about crossing constraints?

- They **delimit the formal language class** of human language (Chomsky & Schutzenberger, 1963).
Why do we care about crossing constraints?

• They **delimit the formal language class** of human language (Chomsky & Schutzenberger, 1963).

• Human language is **not context free** (Shieber, 1985), but also **not fully context-sensitive**.
Why do we care about crossing constraints?

- They **delimit the formal language class** of human language (Chomsky & Schutzenberger, 1963).

- Human language is **not context free** (Shieber, 1985), but also **not fully context-sensitive**.
Why do we care about crossing constraints?

• They **delimit the formal language class** of human language (Chomsky & Schutzenberger, 1963).

• Human language is **not context free** (Shieber, 1985), but also not **fully context-sensitive**.

• It is **mildly context-sensitive** (Weir, 1988; Joshi et al., 1991).
Why do we care about crossing constraints?

- They **delimit the formal language class** of human language (Chomsky & Schutzenberger, 1963).
- Human language is **not context free** (Shieber, 1985), but also **not fully context-sensitive**.
- It is **mildly context-sensitive** (Weir, 1988; Joshi et al., 1991).
- The **mildly context-sensitive languages** are defined by **bounds on gap degree**.
Why do we care about crossing constraints?
Why do we care about crossing constraints?

• Crossing dependencies are implicated in human processing difficulty (Bach et al., 1986; Vogel et al., 1996; Levy et al., 2012; Yadav et al., 2017).
Why do we care about crossing constraints?

- Crossing dependencies are implicated in human processing difficulty (Bach et al., 1986; Vogel et al., 1996; Levy et al., 2012; Yadav et al., 2017).

- Complex interactions with dependency length and predictability (Levy et al., 2012)
Why do we care about crossing constraints?

- Crossing dependencies are implicated in **human processing difficulty** (Bach et al., 1986; Vogel et al., 1996; Levy et al., 2012; Yadav et al., 2017).
- Complex interactions with dependency length and predictability (Levy et al., 2012)
- Processing factors may be the underlying explanation for the rarity and constraints on crossing dependencies (Bach et al., 1986; Ferrer-i-Cancho, 2006).
Why do we care about crossing constraints?

• Crossing dependencies are implicated in human processing difficulty (Bach et al., 1986; Vogel et al., 1996; Levy et al., 2012; Yadav et al., 2017).
• Complex interactions with dependency length and predictability (Levy et al., 2012)
• Processing factors may be the underlying explanation for the rarity and constraints on crossing dependencies (Bach et al., 1986; Ferrer-i-Cancho, 2006).
Our Question
Our Question

• **Observation**: Crossing dependencies appear to be
  • (1) *rare*, and
  • (2) subject to *crossing constraints* (bounded gap degree, 1-end-point-crossing, etc.)
• **Observation**: Crossing dependencies appear to be
  • (1) rare, and
  • (2) subject to *crossing constraints* (bounded gap degree, 1-end-point-crossing, etc.)

• **Question**: Could it be that the apparent crossing constraints (2) are epiphenomenal, arising as a consequence of the rarity of crossing dependencies (1)?
Our Question
Our Question

Observed distribution of gap degree in a treebank
Our Question

Crossing dependencies occur at a **low rate**

Observed distribution of gap degree in a treebank
Our Question

- Crossing dependencies occur at a low rate
- There is a true constraint on gap degree

Observed distribution of gap degree in a treebank
Our Question

- Crossing dependencies occur at a **low rate**
- There is a **true constraint** on gap degree

**Observed distribution of gap degree in a treebank**
Our Question

? 

Crossing dependencies occur at a **low rate** 

There is a **true constraint** on gap degree 

Observed distribution of gap degree in a treebank
Our Question

? Crosses dependencies occur at a low rate

? There is a true constraint on gap degree

Observed distribution of gap degree in a treebank
Our Question
Our Question

• **Question:** Could it be that the apparent crossing constraints are epiphenomenal, arising as a consequence of the rarity of crossing dependencies?
Our Question

• **Question**: Could it be that the apparent crossing constraints are epiphenomenal, arising as a consequence of the rarity of crossing dependencies?

• **Null Hypothesis**: The *observed distribution of crossing constraints* in treebanks can be fully explained by a *low rate of crossing dependencies* (same null hypothesis as in Gómez-Rodríguez & Ferrer-i-Cancho, 2017)
Our Question

- **Question:** Could it be that the apparent crossing constraints are epiphenomenal, arising as a consequence of the rarity of crossing dependencies?

- **Null Hypothesis:** The observed distribution of crossing constraints in treebanks can be fully explained by a low rate of crossing dependencies (same null hypothesis as in Gómez-Rodríguez & Ferrer-i-Cancho, 2017)

- **True Constraint Hypothesis:** It is necessary to posit some additional pressure to explain the observed crossing constraints.
Our Question

• **Question:** Could it be that the apparent crossing constraints are epiphenomenal, arising as a consequence of the rarity of crossing dependencies?

• **Null Hypothesis:** The observed distribution of crossing constraints in treebanks can be fully explained by a low rate of crossing dependencies (same null hypothesis as in Gómez-Rodríguez & Ferrer-i-Cancho, 2017)

• **True Constraint Hypothesis:** It is necessary to posit some additional pressure to explain the observed crossing constraints.

• **Note:** In this work we do not address potential deeper explanations for the low rate of crossings (e.g., dependency length minimization: Ferrer-i-Cancho, 2006)
Our Question

- **Question**: Could it be that the apparent crossing constraints are epiphenomenal, arising as a consequence of the rarity of crossing dependencies?

- **Null Hypothesis**: The observed distribution of crossing constraints in treebanks can be fully explained by a low rate of crossing dependencies (same null hypothesis as in Gómez-Rodríguez & Ferrer-i-Cancho, 2017)

- **True Constraint Hypothesis**: It is necessary to posit some additional pressure to explain the observed crossing constraints.

- **Note**: In this work we do not address potential deeper explanations for the low rate of crossings (e.g., dependency length minimization: Ferrer-i-Cancho, 2006)

- We only ask if a low rate of crossings is sufficient to explain the formal crossing constraints.
Are crossing constraints epiphenomenal?

- Introduction
- Methodology & Baselines
- Results
- Conclusion
Methodology
We implement the null hypothesis as randomly-generated trees with the same rate of crossing dependencies as real trees from UD treebanks.
Methodology

- We implement the null hypothesis as randomly-generated trees with the same rate of crossing dependencies as real trees from UD treebanks.

- Then we test if the crossing constraints are violated at different rates in the real vs. random trees.
Methodology

• We implement the **null hypothesis** as **randomly-generated trees** with the same **rate of crossing dependencies** as real trees from UD treebanks.

• Then we test **if the crossing constraints are violated at different rates** in the real vs. random trees.

• **Random trees**: Uniform random trees generated using Prüfer codes, with the same distribution over sentence lengths as real trees.
Baseline Generation
Baseline Generation

• We control for the rate of crossings using rejection sampling.
Baseline Generation

• We control for the rate of crossings using rejection sampling.
• For each real tree $t$ for a sentence of length $n$ in a treebank,
Baseline Generation

• We control for the rate of crossings using rejection sampling.
• For each real tree $t$ for a sentence of length $n$ in a treebank,
  • **Random tree baseline**: Randomly generate trees of length $n$ until getting one with the same number of crossings as $t$. 
Baseline Generation

• We control for the rate of crossings using rejection sampling.

• For each real tree $t$ for a sentence of length $n$ in a treebank,
  • Random tree baseline: Randomly generate trees of length $n$ until getting one with the same number of crossings as $t$.

• This process is slow, because most random trees and random orders have many more crossings than real trees, especially for longer sentences.
Baseline Generation

• We control for the rate of crossings using rejection sampling.

• For each real tree $t$ for a sentence of length $n$ in a treebank,
  • **Random tree baseline**: Randomly generate trees of length $n$ until getting one with the same number of crossings as $t$.
  • This process is slow, because most random trees and random orders have many more crossings than real trees, especially for longer sentences.
  • So we only do this for trees of length $n \leq 12$. 
Baseline Generation

• We control for the rate of crossings using rejection sampling.
• For each real tree $t$ for a sentence of length $n$ in a treebank,
  • **Random tree baseline**: Randomly generate trees of length $n$ until getting one with the same number of crossings as $t$.
  • This process is slow, because most random trees and random orders have many more crossings than real trees, especially for longer sentences.
  • So we only do this for trees of length $n \leq 12$.
• In order to get a large sample, we combine trees from multiple languages into one large dataset.
Baseline Generation

• We control for the rate of crossings using rejection sampling.

• For each real tree $t$ for a sentence of length $n$ in a treebank,
  • **Random tree baseline**: Randomly generate trees of length $n$ until getting one with the same number of crossings as $t$.

• This process is slow, because most random trees and random orders have many more crossings than real trees, especially for longer sentences.

  • So we only do this for trees of length $n \leq 12$.

• In order to get a large sample, we combine trees from multiple languages into one large dataset.
Crossing Constraints
Crossing Constraints

• The **gap degree** of a node X is the maximum number of discontinuities in chains of dependents emanating from X.
Crossing Constraints

- The **gap degree** of a node $X$ is the maximum number of discontinuities in chains of dependents emanating from $X$. 
Crossing Constraints

- The **gap degree** of a node $X$ is the maximum number of discontinuities in chains of dependents emanating from $X$. 

![Diagram showing gap degree 2 with nodes $X_g$, $X_k$, $X_d$, $X_i$, $X_h$, and $X_j$.]
Crossing Constraints

- The **gap degree** of a node $X$ is the maximum number of discontinuities in chains of dependents emanating from $X$.

- The **edge degree** of arc $X_h \rightarrow X_d$ is the number of nodes between $X_h$ and $X_d$ that are not transitively dominated by $X_h$ (call these “intervening nodes”).
Crossing Constraints

• The **gap degree** of a node \( X \) is the maximum number of discontinuities in chains of dependents emanating from \( X \).

\[
\begin{array}{c}
X_g & \rightarrow & X_k & \rightarrow & X_d & \rightarrow & X_i & \rightarrow & X_h & \rightarrow & X_j \\
\end{array}
\]

Gap degree 2

• The **edge degree** of arc \( X_h \rightarrow X_d \) is the number of nodes between \( X_h \) and \( X_d \) that are not transitively dominated by \( X_h \) (call these “intervening nodes”).

• The number of **end-point crossings** is the number of heads which dominate the intervening nodes between \( X_h \rightarrow X_d \).
Crossing Constraints

• The **gap degree** of a node $X$ is the maximum number of discontinuities in chains of dependents emanating from $X$.

  \[
  \begin{array}{cccccc}
  X_g & X_k & X_d & X_i & X_h & X_j \\
  \end{array}
  \]

• The **edge degree** of arc $X_h \rightarrow X_d$ is the number of nodes between $X_h$ and $X_d$ that are not transitively dominated by $X_h$ (call these “intervening nodes”).

• The number of **end-point crossings** is the number of heads which dominate the intervening nodes between $X_h \rightarrow X_d$.

  \[
  \begin{array}{cccccccc}
  X_d & X_i & X_a & X_b & X_h & X_j & X_r \\
  \end{array}
  \]

(a) : Edge degree=2, End-point crossing=1
Crossing Constraints

• The **gap degree** of a node $X$ is the maximum number of discontinuities in chains of dependents emanating from $X$.

• The **edge degree** of arc $X_h \rightarrow X_d$ is the number of nodes between $X_h$ and $X_d$ that are not transitively dominated by $X_h$ (call these “intervening nodes”).

• The number of **end-point crossings** is the number of heads which dominate the intervening nodes between $X_h \rightarrow X_d$.

(a) : Edge degree=2, End-point crossing=1

(b) : Edge degree=2, End-point crossing=2
Crossing Constraints
Crossing Constraints

• For an arc $X_h \rightarrow X_d$ with an intervener, the **heads’ depth difference** is the difference between the depth of $X_h$ and the head of the intervener.
Crossing Constraints

- For an arc $X_h \rightarrow X_d$ with an intervener, the **heads’ depth difference** is the difference between the depth of $X_h$ and the head of the intervener.
Crossing Constraints

- For an arc $X_h \rightarrow X_d$ with an intervener, the **heads’ depth difference** is the difference between the depth of $X_h$ and the head of the intervener.
Crossing Constraints

• For an arc $X_h \rightarrow X_d$ with an intervener, the **heads’ depth difference** is the difference between the depth of $X_h$ and the head of the intervener.

• HDD is implicated in human processing difficulty (Phillips et al., 2005; Yadav et al., 2017)
Comparing the Real and Random Trees
Comparing the Real and Random Trees

- We test whether crossing constraint violations occur at significantly different rates in real vs. random trees (similar to Courtin & Yan, 2019), as a function of
Comparing the Real and Random Trees

• We test whether **crossing constraint violations** occur at **significantly different rates** in real vs. random trees (similar to Courtin & Yan, 2019), as a function of

  • Sentence length
Comparing the Real and Random Trees

• We test whether crossing constraint violations occur at significantly different rates in real vs. random trees (similar to Courtin & Yan, 2019), as a function of
  • Sentence length
  • Tree depth
Comparing the Real and Random Trees

• We test whether **crossing constraint violations** occur at **significantly different rates** in real vs. random trees (similar to Courtin & Yan, 2019), as a function of
  • Sentence length
  • Tree depth
  • Tree arity
Comparing the Real and Random Trees

- We test whether **crossing constraint violations** occur at **significantly different rates** in real vs. random trees (similar to Courtin & Yan, 2019), as a function of
  - Sentence length
  - Tree depth
  - Tree arity
- To do this, we fit mixed-effects Poisson regressions to predict rates of crossing constraint violations.
Comparing the Real and Random Trees

• We test whether crossing constraint violations occur at significantly different rates in real vs. random trees (similar to Courtin & Yan, 2019), as a function of
  • Sentence length
  • Tree depth
  • Tree arity
• To do this, we fit mixed-effects Poisson regressions to predict rates of crossing constraint violations.
Comparing the Real and Random Trees
Comparing the Real and Random Trees

• For example, to test if gap degree is different in real vs. random trees as a function of sentence length, we set up a Poisson regression:
Comparing the Real and Random Trees

• For example, to test if gap degree is different in real vs. random trees as a function of sentence length, we set up a Poisson regression:

$$\log E[g_i] = \beta_0 + \beta_l|s_i| + \beta_r r_i + \beta_{lr} r_i|s_i| + \gamma_j + \varepsilon,$$
Comparing the Real and Random Trees

• For example, to test if gap degree is different in real vs. random trees as a function of sentence length, we set up a Poisson regression:

\[ \log E[g_i] = \beta_0 + \beta_l|s_i| + \beta_r r_i + \beta_{lr} r_i|s_i| + \gamma_j + \varepsilon, \]

• Where \( g_i \) is the gap degree of the i’th sentence,
Comparing the Real and Random Trees

• For example, to test if gap degree is different in real vs. random trees as a function of sentence length, we set up a Poisson regression:

\[
\log E[g_i] = \beta_0 + \beta_l|s_i| + \beta_r r_i + \beta_{lr} r_i|s_i| + \gamma_j + \epsilon,
\]

• Where \(g_i\) is the gap degree of the i’th sentence,
• \(|s_i|\) is the length of the i’th sentence,
Comparing the Real and Random Trees

• For example, to test if **gap degree** is different in real vs. random trees as a function of sentence length, we set up a Poisson regression:

\[
\log E[g_i] = \beta_0 + \beta_l |s_i| + \beta_r r_i + \beta_{lr} r_i |s_i| + \gamma_j + \epsilon,
\]

- Where \( g_i \) is the gap degree of the \( i \)’th sentence,
- \( |s_i| \) is the length of the \( i \)’th sentence,
- \( r_i \) is an indicator variable for whether the tree is real (1) or random (0),
Comparing the Real and Random Trees

• For example, to test if gap degree is different in real vs. random trees as a function of sentence length, we set up a Poisson regression:

\[
\log E[g_i] = \beta_0 + \beta_l |s_i| + \beta_r r_i + \beta_{lr} r_i |s_i| + \gamma_j + \epsilon,
\]

• Where \(g_i\) is the gap degree of the i’th sentence,
• \(|s_i|\) is the length of the i’th sentence,
• \(r_i\) is an indicator variable for whether the tree is real (1) or random (0),
• \(\gamma_j\) is a random intercept for the j’th language.
Comparing the Real and Random Trees

• For example, to test if gap degree is different in real vs. random trees as a function of sentence length, we set up a Poisson regression:

\[
\log E[g_i] = \beta_0 + \beta_l |s_i| + \beta_r r_i + \beta_{lr} r_i |s_i| + \gamma_j + \epsilon,
\]

• Where \( g_i \) is the gap degree of the \( i \)'th sentence,
• \(|s_i|\) is the length of the \( i \)'th sentence,
• \( r_i \) is an indicator variable for whether the tree is real (1) or random (0),
• \( \gamma_j \) is a random intercept for the \( j \)'th language.

• The beta and gamma parameters are fit to the data.
Comparing the Real and Random Trees

• For example, to test if gap degree is different in real vs. random trees as a function of sentence length, we set up a Poisson regression:

\[ \log E[g_i] = \beta_0 + \beta_l|s_i| + \beta_r r_i + \beta_{lr}r_i|s_i| + \gamma_j + \epsilon, \]

• Where \( g_i \) is the gap degree of the \( i \)’th sentence,
• \( |s_i| \) is the length of the \( i \)’th sentence,
• \( r_i \) is an indicator variable for whether the tree is real (1) or random (0),
• \( \gamma_j \) is a random intercept for the \( j \)’th language.

• The beta and gamma parameters are fit to the data.
• The important coefficient is \( \beta_{lr} \), the interaction coefficient:
Comparing the Real and Random Trees

• For example, to test if gap degree is different in real vs. random trees as a function of sentence length, we set up a Poisson regression:

\[
\log E[g_i] = \beta_0 + \beta_l |s_i| + \beta_r r_i + \beta_{lr} r_i |s_i| + \gamma_j + \varepsilon,
\]

• Where \( g_i \) is the gap degree of the i’th sentence,
• \( |s_i| \) is the length of the i’th sentence,
• \( r_i \) is an indicator variable for whether the tree is real (1) or random (0),
• \( \gamma_j \) is a random intercept for the j’th language.

• The beta and gamma parameters are fit to the data.
• The important coefficient is \( \beta_{lr} \), the interaction coefficient:
  • If it is negative, that means gap degree grows slower with sentence length in real vs. random trees.
Data

• We test on UD v2.3 treebanks of 14 languages:
  • German, English, Hindi, French, Arabic, Russian, Czech, Italian, Spanish, Afrikaans, Japanese, Korean, Bulgarian, Slovak
• We exclude all root and punctuation dependencies.
• We combine trees from all treebanks (but control for language in our regression models).
Are crossing constraints epiphenomenal?

• Introduction
• Methodology & Baselines
• Results
• Conclusion
Gap Degree
Gap Degree

- As a function of **sentence length:**
Gap Degree

- As a function of **sentence length:**

![Graph showing observed and random baseline gap degree as a function of sentence length.](image-url)
Gap Degree

• As a function of **sentence length:**
Gap Degree

• As a function of **sentence length:**

\[ n.s. \]
Gap Degree

- As a function of **sentence length**: 

![Graph showing gap degree as a function of sentence length](image)

- As a function of **tree depth**: 

**n.s.**
Gap Degree

- As a function of **sentence length**:

- As a function of **tree depth**:

  $n.s.$
Gap Degree

• As a function of **sentence length**:

\[
\text{Gap Degree}
\]

\[
\begin{array}{c}
\text{Observed} \\
\text{Random baseline}
\end{array}
\]

\[
\begin{array}{c}
\text{Gap degree} \\
\text{Sentence length}
\end{array}
\]

\[
\begin{array}{c}
2.5 \quad 5.0 \quad 7.5 \quad 10.0 \\
2.5 \quad 5.0 \quad 7.5 \quad 10.0
\end{array}
\]

\[
\begin{array}{c}
n.s.
\end{array}
\]

• As a function of **tree depth**:

\[
\text{Gap Degree}
\]

\[
\begin{array}{c}
\text{Observed} \\
\text{Random baseline}
\end{array}
\]

\[
\begin{array}{c}
\text{Gap degree} \\
\text{Tree depth}
\end{array}
\]

\[
\begin{array}{c}
4 \quad 6 \quad 8 \quad 10 \\
4 \quad 6 \quad 8 \quad 10
\end{array}
\]

\[
\begin{array}{c}
p < .001
\end{array}
\]
End-point Crossings
End-point Crossings

![Graph showing observed and random baseline of end-point crossings vs sentence length]
End-point Crossings

$p < .001$
End-point Crossings

\[ p < .001 \]
End-point Crossings

$p < .001$

$p < .001$
Evidence for the True Constraint Hypothesis?

<table>
<thead>
<tr>
<th>as a function of...</th>
<th>Gap degree</th>
<th>Edge Degree</th>
<th>End-point Crossings</th>
<th>Heads’ Depth Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>~ length</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>~ arity</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>~ depth</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

✓ = significant interaction coefficient
✗ = nonsignificant interaction coefficient
Discussion
Discussion

• **Edge degree** is most distinctively different between real and random trees.
Discussion

• **Edge degree** is most distinctively different between real and random trees.

• **Gap degree** is the *least* distinctively different.
Discussion

- **Edge degree** is most distinctively different between real and random trees.
- **Gap degree** is the *least* distinctively different.
- Most crossing constraints differ between real and random trees *as a function of tree depth.*
• **Edge degree** is most distinctively different between real and random trees.

• **Gap degree** is the *least* distinctively different.

• Most crossing constraints differ between real and random trees *as a function of tree depth*.

• Future work: Control for tree depth, arity, etc. in the random trees.
Are crossing constraints epiphenomenal?

- Introduction
- Methodology & Baselines
- Results
- Conclusion
Conclusion
Conclusion

• Despite 30 years of linguistic formalisms based on a bound on gap degree, gap degree is the constraint for which we have the least evidence for the True Constraint Hypothesis.
Conclusion

- Despite 30 years of linguistic formalisms based on a bound on gap degree, **gap degree is the constraint for which we have the least evidence for the True Constraint Hypothesis.**

- Crossing constraints are most distinctive **as a function of depth**, suggesting a connection to **theories of human sentence processing difficulty based on tree depth** (Yngve, 1960; Gibson & Thomas, 1999; Jing & Liu, 2015; Komori et al., 2019).
Conclusion

• Despite 30 years of linguistic formalisms based on a bound on gap degree, **gap degree is the constraint for which we have the least evidence for the True Constraint Hypothesis.**

• Crossing constraints are most distinctive **as a function of depth**, suggesting a connection to **theories of human sentence processing difficulty based on tree depth** (Yngve, 1960; Gibson & Thomas, 1999; Jing & Liu, 2015; Komori et al., 2019).

• Future work can control for other factors:
Conclusion

• Despite 30 years of linguistic formalisms based on a bound on gap degree, **gap degree is the constraint for which we have the least evidence for the True Constraint Hypothesis.**

• Crossing constraints are most distinctive **as a function of depth**, suggesting a connection to **theories of human sentence processing difficulty based on tree depth** (Yngve, 1960; Gibson & Thomas, 1999; Jing & Liu, 2015; Komori et al., 2019).

• Future work can control for other factors:
  • Tree depth and arity
Conclusion

• Despite 30 years of linguistic formalisms based on a bound on gap degree, **gap degree is the constraint for which we have the least evidence for the True Constraint Hypothesis.**

• Crossing constraints are most distinctive **as a function of depth**, suggesting a connection to **theories of human sentence processing difficulty based on tree depth** (Yngve, 1960; Gibson & Thomas, 1999; Jing & Liu, 2015; Komori et al., 2019).

• Future work can control for other factors:
  • Tree depth and arity
  • Dependency length
Conclusion

• Despite 30 years of linguistic formalisms based on a bound on gap degree, gap degree is the constraint for which we have the least evidence for the True Constraint Hypothesis.

• Crossing constraints are most distinctive as a function of depth, suggesting a connection to theories of human sentence processing difficulty based on tree depth (Yngve, 1960; Gibson & Thomas, 1999; Jing & Liu, 2015; Komori et al., 2019).

• Future work can control for other factors:
  • Tree depth and arity
  • Dependency length
  • Could controlling crossing constraints explain the rarity of crossings?
Thanks all!

- All code is available online at https://github.com/yadavhimanshu059/measures_of_nonProjectivity

- Thanks to Roger Levy and Tim O’Donnell for discussion, and to our SyntaxFest reviewers for helpful suggestions.

- Thanks to the TLT organizers!