From the world to word order: Deriving biases in noun phrase order from statistical properties of the world

Jennifer Culbertson, Marieke Schouwstra, Simon Kirby
Centre for Language Evolution
University of Edinburgh
Why are languages the way they are?

Language variation is constrained in specific ways. What are the origins of these constraints?
- How specific are they to language?
- To what extent are they learned, or innate?

linguistic system
cognitive system
history
## Noun phrase word order

<table>
<thead>
<tr>
<th>Language</th>
<th>Noun phrase</th>
<th>Word Order</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English</strong></td>
<td>Those two purple chairs</td>
<td>Dem Num Adj N</td>
</tr>
<tr>
<td><strong>Thai</strong></td>
<td>เก้าอี้ม่วงสองพวกนี้</td>
<td>N Adj Num Dem</td>
</tr>
<tr>
<td><strong>French</strong></td>
<td>Ces deux chaises violettes</td>
<td>Dem Num N Adj</td>
</tr>
<tr>
<td><strong>Indonesian</strong></td>
<td>Dua kursi ungu itu</td>
<td>Num N Adj Dem</td>
</tr>
<tr>
<td><strong>Kîîtharakka</strong></td>
<td>Mabuku mara mathatu manene</td>
<td>N Dem Num Adj</td>
</tr>
</tbody>
</table>
Noun phrase word order

(Dryer 2018)
Hierarchical structure of noun phrase

[Those [two [purple chairs]]]

Dem  Num  Adj  N
Homomorphic orders
Homomorphic orders
Research questions

Q1: Is there a cognitive bias for homomorphic mappings between underlying structure and linear order?
Research questions

Q2: What is the origin of the underlying structure?
Outline

1. Is there a bias for homomorphism?
   - Previous experimental evidence from artificial language learning
   - New experimental evidence from silent gesture

2. What is the origin of the underlying structure?
   - Conceptual representation reflecting strength of association
   - Corpus-based evidence for learnability
   - Experimental evidence for connection between strength of association and gesture order
Evidence for homomorphism: ALL

Given ambiguous input, what order do participants infer?

Trained on order of \textbf{N and single modifier}... Tested on phrases with \textbf{multiple modifiers}

Evidence for homomorphism: ALL

Prediction based on **homomorphism**:

(Culbertson & Adger 2014, *PNAS*; Martin et al. to appear, *Glossa*)
Evidence for homomorphism: ALL

Prediction based on **English surface order:** 

\[ \text{N Dem Num Adj} \]

--

(Culbertson & Adger 2014, *PNAS*; Martin et al. to appear, *Glossa*)

Language Meet the Authors Webinar, August 2020
Training

éyèy tákà

feather red

(Martin et al. to appear, Glossa)
Testing

feather red that

or

feather that red
Results

(Martin et al. to appear, Glossa)
Interim summary

• Participants were more likely to infer Adj closest to N, Dem farthest away
• Could there nevertheless be transfer from English here (e.g. mirror image order)?
• Need a methodology where transfer is minimised.

Silent gesture paradigm

• Participants must use (only) their hands to convey meaning
• Less influence of conventionalized spoken language on behavior (e.g., Goldin-Meadow et al. 2008)
Silent gesture and basic word order

Different L1 backgrounds, different spoken word order

---

![Bar graph showing the proportion of SOV order in speech for different language groups.](image)
Silent gesture and basic word order

Given **NO input**, what gesture order do participants use?

- **GIRL**
- **BOX**
- **COVERS**

- **CAPTAIN**
- **PAIL**
- **SWINGS**
Silent gesture and basic word order

Different L1 backgrounds, same gesture order

![Graph showing the proportion of strings consistent with ArPA order for speech and gesture in different languages.](image)

- **Speech**:
  - Turkish speakers: 0.94 (SE 0.000001)
  - Chinese speakers: 0.90 (SE 0.000001)
  - English speakers: 0.85 (SE 0.000001)
  - Spanish speakers: 0.83 (SE 0.000001)

- **Gesture**:
  - Turkish speakers: 1.00
  - Chinese speakers: 0.98 (SE 0.000001)
  - English speakers: 0.97 (SE 0.000001)
  - Spanish speakers: 0.96 (SE 0.000001)
Gesture order is sensitive to **conceptual structure**
Gesture order is sensitive to **conceptual structure**

Schouwstra & de Swart (2014)
Evidence for homomorphism: Silent gesture

Given NO input, what order do participants use?
Stimuli
Example gesture sequences

these four spotty squares

these four squares spotty
Example gesture sequences

those four stripy triangles

those triangles stripy four
Results: individual modifier orders

% Post-nominal

- Adj
- Num
- Dem

Modifier type
Results: raw frequency of gesture orders
Results: by-participant preferred patterns
Interim summary

English speakers show a homomorphism bias

• Participants infer Adj closest to N, Dem farthest away
• They also improvise these orders in gesture

But where does this underlying structure come from?

• Is it about syntactic structure or conceptual structure?
• Is it innate or learned?
An intuition...
A hypothesis:

The underlying structure of the NP is based on conceptual structure, which encodes strength of association:

- properties
- numerosities
- location/discourse status

This is learnable from observing the world.

How do we evaluate this?

- An information-theoretic measure of strength of association
- Corpus as a proxy for the world
Evaluating strength of association

Point-wise mutual information:

\[ pmi(x, y) = \log \left( \frac{p(x, y)}{p(x)p(y)} \right) \]

Is the frequency of the pair higher than expected given the individual frequencies of the words alone?
Evaluating strength of association

Point-wise mutual information:

Is the frequency of the pair higher than expected given the individual frequencies of the words alone?

PMI in a corpus

1. Get all (N, Mod) pairs and their frequencies
2. Discard very low frequency pairs
3. Calculate pmi for each pair
4. Average pmi for each modifier type
Results: English

High pmi:
- alcoholic beverage
- dense vegetation
- seven founders

Low pmi:
- that child
- new fact
- one program
Results: cross-linguistic validation

![Graph showing the results of cross-linguistic validation for various languages.](image)
Interim summary

Strength of association differs among modifiers

\{\text{objects, properties}\} > \{\text{objects, numerosities}\}

\{\text{objects, numerosities}\} > \{\text{objects, discourse status}\}

- observe objects in the world
- use to build representation of conceptual structure
- map to syntactic structure
Interim summary

Strength of association differs among modifiers

\{\text{objects, properties}\} > \{\text{objects, numerosities}\}

\{\text{objects, numerosities}\} > \{\text{objects, discourse status}\}

Mediated by syntactic categories

Hierarchy is \textit{not} constructed on a phrase-by-phrase or modifier-by-modifier basis...

- map to syntactic structure
Adjective categories and PMI


[Big [purple chairs]]

Size Color N
Adjective categories and PMI

Q3: Can we see the effect of adjective PMI on gesture order?
Adjective categories and PMI

Q3: If Adj has lower PMI then will, e.g., N-Num-Adj gesture order be more likely?
Strength of association and order

Texture condition

Size condition
Example gesture sequences

*those four stripy toothbrushes*

*those toothbrushes stripy four*
Example gesture sequences

*these four spotty pencils*

*pencils spotty four these*
Example gesture sequences

*those five small toothbrushes*

*those toothbrushes five small*
Results: raw frequency of gesture orders

Texture

Frequency

Size

Frequency

Language Meet the Authors Webinar, August 2020
Interim summary

Manipulating Adj PMI changes the strength of the bias
• Participants in the lower PMI size condition produce more gestures with non-homomorphic order of Adj and Num

But the homomorphism bias is strong...
• Lower PMI adjectives are still very likely to be gestured in closer proximity to the Noun
Conclusion

Universal constraints on NP word order

- Underlying structure: a learnable conceptual asymmetry (evidence: corpus study)
- Homomorphism: a bias for transparent mappings (evidence: experiments)
A (more) complete picture

- observe objects in the world
- use to build representation of conceptual structure
- map to syntactic structure
- + homomorphism bias = common linear orders

A (more) complete picture

learned

language specific

language specific

- observe objects in the world
- use to build representation of conceptual structure
- map to syntactic structure
- + homomorphism bias = common linear orders

A (more) complete picture

learned

language specific

language specific
A (more) complete picture

- observe objects in the world
- use to build representation of conceptual structure
- map to syntactic structure
- + homomorphism bias = common linear orders

Learned language specific
From the world to word order: Deriving biases in noun phrase order from statistical properties of the world

Jennifer Culbertson, Marieke Schouwstra, Simon Kirby
Centre for Language Evolution
University of Edinburgh
Shape of the distribution

- Frequency according to Cysouw (2010)
  - Raw language numbers

- Frequency according to Dryer (2018)
  - Adjusted for genetic relations
  - By genus
  - Raw language numbers