Structural Frequency Effects in Noisy-channel Comprehension

Yingtong Liu¹,³, Rachel Ryskin², Richard Futrell⁴, Edward Gibson³

¹ Harvard University
² University of California, Merced
³ Massachusetts Institute of Technology
⁴ University of California, Irvine
Communication in the context of noise

• Noise is present in typical language use.
• Comprehenders can often successfully figure out what the speaker wants to convey, even when the utterance is corrupted by noise.

uhh..you have to ..uh.. square it facely

You have to face it squarely.
Comprehension of possibly corrupted input

Sentence: The mother gave the candle the daughter.
Question: Did the daughter receive something/someone?

- Yes (Non-literal response) 52%
- No (Literal response) 48%

(Gibson et al., 2013)

• How to capture these observations?
The noisy-channel model of sentence comprehension

- The language comprehension mechanism is well-designed for recovering the intended utterance from noisy input.

\[
P(S_{\text{intended}} \mid S_{\text{perceived}}) \propto P(S_i) \cdot P(S_i \rightarrow S_p)
\]

Probability of the intended sentence given the perceived sentence

The noise model

Prior probability based on world and linguistic knowledge

(e.g., Levy, 2008; Gibson et al., 2013; Poppels & Levy, 2016; Ryskin et al., 2018; Keshev & Meltzer-Asscher, 2021)
Sentence comprehension over a noisy-channel

Sentence: The mother gave the candle the daughter.
Question: Did the daughter receive something/someone?

✓ Yes (Non-literal response)
〇 No (Literal response)

(Gibson et al., 2013)
Sentence comprehension over a noisy-channel

\[ P(S_i | S_p) \propto P(S_i) \, P(S_i \rightarrow S_p) \]

\[ P(S_p | S_p) \propto P(S_p) \, P(S_p \rightarrow S_p) \]

\[
\begin{array}{ccc}
\text{medium} & \text{low} & \text{high} \\
\text{high} & \text{high} & \text{medium}
\end{array}
\]
Our proposal

• Previous works have mainly focused on the meaning prior
  \[ P(S_i \mid S_p) \propto P(S_i) \ P(S_i \rightarrow S_p) \]

• \[ P(s_i) = P(s_{i\_structure}, s_{i\_meaning}) \]
  \[ P(S_i \mid S_p) \propto P(s_{i\_structure}, s_{i\_meaning}) \ P(S_i \rightarrow S_p) \]
  (cf. Bergen et al., 2012; Poppels & Levy, 2016; Keshev & Meltzer-Asscher, 2021)

• Prediction of our proposed noisy-channel model
  ➢ Comprehenders draw more inferences for sentences formed in low-frequency structures compared to those formed in high-frequency structures.
Experiments 1-2 & Corpus search

• **Goal:** Quantitatively measure the degree to which English and Mandarin Chinese allow the six logically possible word orders (SVO, OSV, SOV, VOS, OVS, VSO).

  - While there is consensus about English, it is unclear to what extent Chinese allows various word orders and there has been no experimental measurement for that.
Experiments 1-2

- Behavioral experiments in English and Chinese (N=30 in both languages):
Results

• SVO is more frequently chosen than OSV, while other word orders are less common ($\beta$s $>$ 8, $p$s $<$ 0.01).

• Mandarin has a higher word order flexibility than English. The entropy in the distribution of allowable Mandarin word orders ($H$ = 1.28, 95% CI = [1.26, 1.3]) is larger than that in English ($H$ = 0.95, 95% CI = [0.9, 1]).
Corpus search

- Frequency of SVO and OSV word order in English and Chinese Penn Treebank.

Results from written text in Penn Treebank

<table>
<thead>
<tr>
<th></th>
<th>Raw counts</th>
<th>Probability of each structure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SVO</td>
<td>OSV</td>
</tr>
<tr>
<td></td>
<td>SVO</td>
<td>OSV</td>
</tr>
<tr>
<td>English</td>
<td>87515</td>
<td>176</td>
</tr>
<tr>
<td></td>
<td>0.62</td>
<td>0.001</td>
</tr>
<tr>
<td>Chinese</td>
<td>65688</td>
<td>1693</td>
</tr>
<tr>
<td></td>
<td>0.64</td>
<td>0.015</td>
</tr>
</tbody>
</table>
Experiments 3-4

- **Goal:** Test our proposed noisy-channel model with consideration of the structural prior
  - English and Chinese speakers’ comprehension of implausible and plausible sentences formed in SVO or OSV.

<table>
<thead>
<tr>
<th>Sample trial (English)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sentence:</strong> The trash threw the boy. <em>(SVO_implausible)</em></td>
</tr>
<tr>
<td><strong>Question:</strong> Did the boy throw something/someone?</td>
</tr>
<tr>
<td>☐ Yes <em>(Noisy-channel inference response)</em></td>
</tr>
<tr>
<td>☐ No <em>(Literal response)</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Example materials (English):</th>
</tr>
</thead>
<tbody>
<tr>
<td>The boy threw the trash. <em>(SVO_plausible)</em></td>
</tr>
<tr>
<td>The trash, the boy threw. <em>(OSV_plausible)</em></td>
</tr>
<tr>
<td>The trash threw the boy. <em>(SVO_implausible)</em></td>
</tr>
<tr>
<td>The boy, the trash threw. <em>(OSV_implausible)</em></td>
</tr>
</tbody>
</table>

(N=97 for English; N=81 for Chinese).
Results

- OSV sentences were more likely to be interpreted non-literally compared to SVO sentences in both English and Mandarin (βs>1.5, zs>6.5, ps<0.01).

- As predicted, people are much more likely to interpret “The boy, the trash threw” as the more plausible “The boy threw the trash” than they are to interpret “The trash threw the boy” in the more plausible way.

→ Comprehenders draw more inferences for low-frequency constructions, supporting our proposed noisy-channel model with implementation of the structure prior.
What kind of frequencies are comprehenders tracking?

- The ‘grain sizes’ distributional syntactic information stored by language users.

<table>
<thead>
<tr>
<th>Pair</th>
<th>Sentence</th>
<th>Construction</th>
<th>String</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple transitives</td>
<td>The trash threw the boy.</td>
<td>SVO</td>
<td>NVN</td>
</tr>
<tr>
<td></td>
<td>The boy, the trash threw.</td>
<td>OSV</td>
<td>NNV</td>
</tr>
<tr>
<td>Clefts</td>
<td>It was the trash that threw the boy.</td>
<td>Subj cleft</td>
<td>NVN</td>
</tr>
<tr>
<td></td>
<td>It was the boy that the trash threw.</td>
<td>Obj cleft</td>
<td>NNV</td>
</tr>
</tbody>
</table>

**Construction-based hypothesis:**
Comprehenders track frequencies of each of the four constructions (Goldberg, 2016; Abeillé et al., 2020).

**Linear string-based hypothesis:**
The comprehension mechanism only tracks two kinds of strings – NVN for SVO & subj cleft, and NNV for OSV and obj cleft. (Bates et al., 1982; Ferreira 2003).

**Prediction:** Variation between SVO/OSV and clefts.

**Prediction:** NO variation between SVO/OSV and clefts.
Experiment 5

• Goals:

➢ Test of our proposed noisy-channel model - subject/object clefts.

➢ Evaluation of the two hypotheses about distributional syntactic information stored by language users (construction-based vs. linear string-based) with SVO/OSV and clefts.

⇒ Expt 5a: replication of Expt 3 of English SVO/OSV with a new paradigm
⇒ Expt 5b: test of English clefts
Experiments 5

• Paradigm:

_ _ _ _ _ _ _ _ _ _ _ _ _
It ___ ___ ___ ___ ___ ___ ___ ___
threw ___ ___
boy.
1. Comprehension question about the sentence you just read

Did the boy throw something/someone?
• Yes
• No
2. Please type the sentence exactly in the form you just read.
Prediction of the comprehension task

• Our proposed noisy-channel model:
  ➢ More inferences for low-frequency constructions

➢ Is there difference between clefts and SVO/OSV?
**Construction-based hypothesis**: the difference in inference rate is larger between SVO and OSV than between the two types of clefts (SVO: OSV = 497: 1; Subject vs. Object cleft = 17:1 in Penn Treebank)

**Linear string-based hypothesis**: the inference rate difference between SVO and OSV should be similar to that between subj and obj clefts.
Results of the comprehension questions

- Replicated Expt3: OSV sentences were more likely to be interpreted non-literally than SVO ($p<0.01$).
- As predicted, more non-literal responses for object clefts than for subject clefts ($p<0.02$).

→ Comprehenders draw more inferences for low-frequency constructions

- The difference in inference rate between SVO vs. OSV is larger than that between subject and object clefts ($p<0.01$), supporting the construction-based hypothesis, not the linear string-based hypothesis.
Where do the non-literal responses come from?

- Our **speaker’s/transmission channel hypothesis**: Comprehenders are fully aware of the input sentence. They draw rational inferences about the intended utterance possibly corrupted due to the speaker/transmission procedure.

- Alternative **comprehender’s channel hypothesis**: Comprehenders are NOT fully aware of the input sentence. They mis-read/mis-retrieve the input sentence, leading to an ‘incorrect’ interpretation (c.f., Ferreira, 2003).

### Sample trial

**Sentence**: The boy, the trash threw.

**Question**: Did the boy throw something/someone?

- Yes (Non-literal response)
- No (Literal response)
Predictions of the retyping task:

- **Speaker’s channel hypothesis:** (in our proposed noisy-channel model)
  Almost no incorrect retyping

- **Comprehender’s channel hypothesis:**
  The proportion of incorrect retyping aligns the amount of non-literal responses
Results of the retyping task:

- The proportion of non-literal responses do not align the amount of incorrect retyping. No error in retyping even for non-literal responses.

→ No evidence for the comprehender’s channel hypothesis. Our speaker’s channel hypothesis is more on the right track.
Discussion

• Supportive evidence for structural frequency effects in comprehenders’ noisy-channel processing.

• Comprehenders track each construction separately, not just the linear string.

• The non-literal interpretations come from comprehenders’ rational inferences of the intended utterance, rather than misperception.
Thank you!

Questions?
Contact: y_liu@g.harvard.edu

Acknowledgments
Dr. Jesse Snedeker
Dr. Kevin Ryan
Language Lab @ MIT
Snedeker Lab @ Harvard University

Funding
NSF Linguistics 1534318 to EG

Edward Gibson
Rachel Ryskin
Richard Futrell