Certain ungrammaticality or uncertain grammaticality: Deciding between frequent errors and infrequent grammatical structures

Maayan Keshev & Aya Meltzer-Asscher, Tel-Aviv University
Introduction
How do humans recover the intended meaning from noisy utterances?

Production/perception errors are very common in language use.

Recent studies suggest that comprehends engage in (Bayesian) rational noisy-channel processing - interpretation is pulled towards more probable “near-neighbors”, depending on:

- The likelihood of the corruption and of specific errors in the input
- The prior probability of the competing utterances

\[ P(\text{utterance} / \text{perceived input}) \propto P(\text{perceived input} / \text{utterance}) \cdot P(\text{utterance}) \]

Levy 2008; Levy et al., 2009; Gibson et al., 2013
Introduction
How do humans recover the intended meaning from noisy utterances?

Levy, Bicknell, Slattery, & Rayner (2009): Readers regress and reread an early function word when it has near neighbors that would allow a more probable structure.

(1) The coach smiled at the player tossed the frisbee
→ Replace ‘at’ with: ‘as’/’and’ to avoid the reduced relative clause?

Gibson, Bergen, & Piantadosi (2013): Readers endorse a noisy-input analysis of semantically implausible sentences if the required edit is minor.

(2) The mother gave the candle the daughter
→ The mother gave the candle to the daughter
The current study
How improbable should a structure be for readers to consider it noisy?

We test the effect of uncertainty and prior probability using Hebrew relative clauses with temporary ambiguity [object relative vs. subject with an agreement mismatch].

In four self-paced reading experiments and one production study, we show that the interpretation of agreement mismatch is modulated by the prior probabilities of the competing analyses.

Corrupted input, very probable structure

Pristine input, very improbable structure
Experiments 1A&B
The temporary SR/OR ambiguity in Hebrew

pagasnu et ha-studentim še-hikir
We-met ACC the-students.PL that-knows.SG
Experiments 1A&B
The temporary SR/OR ambiguity in Hebrew

pagasnu et ha-studentim še-hikir
We-met ACC the-students.PL that-knows.SG

Corrupted SR
via editing the filler studentim → student
or the verb hikir → hikiru
Experiments 1A&B
The temporary SR/OR ambiguity in Hebrew

pagasnu et ha-studentim še-hikir
We-met ACC the-students.PL that-knows.SG

Corrupted SR
(verb-filler mismatch)

et ha-mazkira
ACC the-secretary
Experiments 1A&B
The temporary SR/OR ambiguity in Hebrew

OR + post-verbal subject (very rare)

pagasnu et ha-studentim še-hikir
We-met ACC the-students.PL that-knows.SG

et ha-mazkira
ACC the-secretary

Corrupted SR (verb-filler mismatch)
Experiments 1A&B
The temporary SR/OR ambiguity in Hebrew

pagasnu et ha-studentim še-hikir
We-met ACC the-students.PL that-knows.SG

Corrupted SR
(verb-filler mismatch)

OR + post-verbal subject (very rare)

ha-dikan
the-Dean.NOM

et ha-mazkira
ACC the-secretary
Experiments 1A&B
The temporary SR/OR ambiguity in Hebrew

Rational noisy-channel inference predicts that a SR will be formed → A post-verbal subject should cause processing difficulty.
Experiment 1A

Materials

We met

\[
\text{BASLINE: because}
\]

by the end of term decided the principal to expel him due to poor grades.

'We met because, by the end of term, the principal decided to expel the student due to poor grades'.
Experiment 1A
Materials

We met

\[
\begin{align*}
\text{BASLINE:} & \quad \text{because} \\
\text{MATCH:} & \quad \text{the student.}_{SG} \quad \text{that} \\
\end{align*}
\]

by the end of term decided._SG eventually \textbf{the principal} to expel him due to poor grades.

'We met the student that, by the end of term, the principal decided to expel due to poor grades'.

Corrupted SR vs. rare OR
Processing disruption at the subject?
SPR, 36 participants, 30 sets + 45 fillers (grammatical)
Experiment 1A
Materials

We met

\[
\begin{align*}
\text{BASLELINE:} & \quad \text{because} \quad \text{the student}\_\text{SG that} \\
\text{MATCH:} & \quad \text{the students}\_\text{PL that} \\
\text{MISMATCH:} & \quad \text{by the end of term decided}\_\text{SG eventually the principal to expel them due to poor grades.}
\end{align*}
\]

'We met the students that, by the end of term, the principal decided to expel due to poor grades'.
Experiment 1A

Results

Increased RTs following the post-verbal subject, in the **match** condition relative to the unambiguous baseline.

→ Readers constructed a **SR**
Experiment 1A

Results

The post-verbal subject reveals increased RTs also in the mismatch condition (relative to baseline).

→ Readers form a SR even when the verb mismatches the filler

Corrupted SR vs. rare OR
Processing disruption at the subject?
SPR, 36 participants, 30 sets + 45 fillers (grammatical)
We met BASLINE: because the student. SG-M that 
MATCH: the students. PL-M that 
1-MISMATCH: the students. PL-F that 
2-MISMATCH: by the end of term decided. SG eventually the principal to expel them due to poor grades.

'We met the students that, by the end of term, the principal decided to expel due to poor grades'.
Experiment 1B

Results

Replication of the findings of Experiment 1a.

Corrupted SR vs. rare OR
Processing disruption at the subject?
SPR, 48 participants, 28 sets + 45 fillers (grammatical)
Experiment 1B

Results

Replication of the findings of Experiment 1a.

Extension to 2-feature mismatch.

Corrupted SR vs. rare OR
Processing disruption at the subject?

SPR, 48 participants, 28 sets + 45 fillers (grammatical)
Intermediate summary

Experiments 1A&B (SPR): Readers assume a corrupted SR, rather than construct an OR with a post-verbal subject.
Intermediate summary

Experiments 1A&B (SPR): Readers assume a corrupted SR, rather than construct an OR with a post-verbal subject.

But plural verbs allow another outlet:

OR + impersonal null subject

<table>
<thead>
<tr>
<th>SR with an agreement mismatch</th>
<th>OR with a post verbal subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrupted input, very probable structure</td>
<td>Pristine input, somewhat probable?</td>
</tr>
</tbody>
</table>
Experiments 2A&B
The temporary SR/OR ambiguity in Hebrew (with a plural verb)

OR + post-verbal subject (very rare)

pagasnu et ha-student še-hikiru
We-met ACC the-student.SG that-knew.PL

ha-dikan
the-Dean.NOM

et ha-mazkira
ACC the-secretary

Corrupted SR
via editing the filler student → studentim
or the verb hikiru → hikir
Experiments 2A&B
The temporary SR/OR ambiguity in Hebrew (with a plural verb)

pagasnu et ha-student še-hikiru
We-met ACC the-student.SG that-knew.PL

Corrupted SR
(verb-filler mismatch)

et ha-mazkira
ACC the-secretary
Experiments 2A&B
The temporary SR/OR ambiguity in Hebrew (with a plural verb)

OR + impersonal null subject
(somewhat probable)

pagasnu et ha-student še-∅-hikiru
We-met ACC the-student.SG that-∅-know.PL

Corrupted SR
(verb-filler mismatch)

et ha-mazkira
ACC the-secretary
Experiments 2A&B
The temporary SR/OR ambiguity in Hebrew (with a plural verb)

OR + impersonal null subject
(somewhat probable)

*oto ba-oniversita*
him in.the-university

*et ha-mazkira*
ACC the-secretary

Corrupted SR
(verb-filler mismatch)

*pagasnu et ha-student še-Ø-hikiru*
We-met ACC the-student.SG that-Ø-know.PL

*et ha-mazkira*
ACC the-secretary
Experiments 2A&B

The temporary SR/OR ambiguity in Hebrew (with a plural verb)

**OR + impersonal null subject**
(somewhat probable)

- **oto ba-oniversita**
  him in.the-university

- **et ha-mazkira**
  ACC the-secretary

**Corrupted SR**
(verb-filler mismatch)

- **pagasnu et ha-student**
  še-Ø-hikiru
  We-met ACC the-student.SG that-Ø-know.PL

**OR (+null subject)** → An object should cause a processing difficulty (“filled-gap” effect).

**SR** → No processing difficulty is expected at the object.
Experiment 2A
Materials

The teacher looked for

\[
\begin{align*}
\text{BASLINE:} & \quad \text{the student} & \text{after we} & \text{found.}_{PL} \\
\end{align*}
\]

eventually \text{ACC the bag} of him at the school’s playground.

'The teacher looked for the student after we eventually found his bag at the school’s playground'.
Experiment 2A

Materials

The teacher looked for

- **BASLINE:** the student after we found.
- **CLASSIC FGE:** the student that we eventually found.PL eventually **ACC the bag** of him at the school’s playground.

'The teacher looked for the student that we eventually found his bag at the school’s playground'.

Corrupted SR vs. a simple OR
Processing disruption at the object?
SPR, 48 participants, 24 sets + 51 fillers (grammatical)
Experiment 2A
Materials

The teacher looked for

\[
\begin{align*}
\text{BASLINE:} & \quad \text{the student after we} \\
\text{CLASSIC FGE:} & \quad \text{the student that we} \\
\text{MATCH:} & \quad \text{the students}_{\text{PL}} \text{ that}
\end{align*}
\]

found\text{PL} eventually ACC the bag of them at the school’s playground.

'The teacher looked for the students that eventually found their bag at the school’s playground'.

Corrupted SR vs. a simple OR
Processing disruption at the object?
SPR, 48 participants, 24 sets + 51 fillers (grammatical)
Experiment 2A
Materials

The teacher looked for

\[
\begin{align*}
\text{BASELINE:} & \quad \text{the student after we} \\
\text{CLASSIC FGE:} & \quad \text{the student that we} \\
\text{MATCH:} & \quad \text{the students.}^\text{PL} \quad \text{that} \\
\text{MISMATCH:} & \quad \text{the student.}^\text{SG} \quad \text{that} \\
\end{align*}
\]

found.\text{PL} \quad \text{eventually ACC the bag of him at the school’s playground.}

'The teacher looked for the student that eventually found his bag at the school’s playground'.
In the **classic FGE** condition, increased RTs at the direct object, relative to baseline.

→ Readers constructed an **OR**

In the **match** condition, no slowdown at the object.

→ Readers constructed a **SR**

**Corrupted SR vs. a simple OR**
Processing disruption at the object?

**Results**

SPR, 48 participants, 24 sets + 51 fillers (grammatical)
Experiment 2A
Results

In the mismatch condition, increased RTs at the direct object, relative to baseline.
(In alignment with that observed in the classic FGE condition).

→ Readers constructed an OR, when the impersonal subject reading is available.

Corrupted SR vs. a simple OR
Processing disruption at the object?
SPR, 48 participants, 24 sets + 51 fillers (grammatical)
Experiment 2B
Materials

The teacher looked for

\[
\begin{align*}
\text{BASLINE:} & \quad \text{the student} \quad \text{after we} \\
\text{MATCH:} & \quad \text{the students}_{\text{PL}} \quad \text{that } \emptyset \\
\text{MISMATCH:} & \quad \text{the student}_{\text{SG}} \quad \text{that } \emptyset
\end{align*}
\]

\text{by the end of the break}, \text{found}_{\text{PL}} \text{eventually} \\
\text{ACC the bag} \text{of him at the school’s playground.}

'The teacher looked for the student that by the end of the break eventually found his bag at the school’s playground.'
Experiment 2B

Results

Replication with extended filler-verb distance.

Corrupted SR vs. a simple OR
Processing disruption at the object?

SPR, 36 participants, 24 sets + 51 fillers (grammatical)
Intermediate summary

Experiments 1A&B (SPR): Readers prefer a corrupted SR reading, over constructing an OR with a post-verbal subject.

Experiments 2A&B (SPR): When an impersonal null subject can be licensed (\textsc{verb.pl}), such an OR is formed.

Experiment 3 aims to replicate both findings within one experiment, and to tap into readers’ preferences more directly (a sentence completion task).
Experiment 3
Design

A sentence completion experiment

RSVP of the preamble - sentences truncated after the verb

Four conditions - crossing number agreement on the verb & the filler


Responses coded as SR/OR completions.

FILLER.PL + VERB.SG $\rightarrow$ Form a SR
FILLER.SG + VERB.PL $\rightarrow$ Form an OR
Production, 100 participants, 24 sets + 12 OR fillers
When the filler and the verb match -

• Strictly SR completions, when an impersonal subject isn’t licensed.
• Mostly SR completions, when an impersonal null subject can be licensed.
Experiment 3
Results

And for filler-verb mismatch -
• Mostly SR completions (ignoring subject-verb agreement), when only a post-verbal can restore grammaticality.
Experiment 3

Results

And for filler-verb mismatch -

• Mostly SR completions (ignoring subject-verb agreement), when only a post-verbal can restore grammaticality.

• Mostly OR completions, when an impersonal subject can be licensed.
Discussion

Our study demonstrates rational noisy channel inference in Hebrew, showing that:

• During online processing, readers **apply elaborate probabilistic knowledge** regarding the distribution of structures in the language.

• They are willing to **compromise subject-verb agreement** to allow for a more probable structure (cf. Frazier, 1987) in some cases, depending on the prior probability of the alternative analyses.

| Corrupted input, very probable structure | Pristine input, somewhat probable? | Pristine input, very improbable structure |
Thank you

Jesse Harris
Psycholinguistics seminar at UCLA
Radan Nasrallah
CUNY reviewers

Contact: maayankeshev@mail.tau.ac.il
The likelihood of the different agreement errors in the current study

Deletions are more likely (Gibson et al, 2013). Does it affect our finding?

**EXP2:**

**EXP1:**

**Impersonal subject – Post-verbal subject**

[VERB.pl] חפש – חיפש[VERB.sg]

[NOUN.sg] המלצר – המלצר [NOUN.pl]

We cannot determine if the deletion/addition contrast is a confound since we don’t know for certain weather readers amend the representation of the verb or a that of the filler.
Deletions are more likely (Gibson et al, 2013). Does it affect our finding?

EXP3:

Impersonal subject – Post-verbal subject

[VERB.sg] חיפש – חיפשו [VERB.pl]
[NOUN.sg] המלצרית – המלצריות [NOUN.pl]

In experiment 3 this problem was removed by using the verbs and fillers in their feminine form. In this case, amending the verb requires substitution in both conditions. Amending the filler is more likely in the impersonal subject condition (in contrast to our prediction, and thus does not confound our findings).