Prosody and Listeners’ Interpretation of Ambiguous Discourse

Joseph Tyler
University of Michigan
jctyler@umich.edu
Did Sally mean that she read about housing prices and watched a cool documentary... in history class? separate from history class?
Two Interpretations

I sat in on a history class. [S1]
I read about housing prices. [S2]
And I watched a cool documentary. [S3]

SUBORD

I sat in on a history class.

I read about housing prices.

And I watched a cool documentary.

S1 event elaborated by S2 and S3

COORD

I sat in on a history class.

I read about housing prices.

And I watched a cool documentary.

S1, S2, and S3 each describe three separate events
Prosody in Discourse Production

• Prosody correlates with discourse structure in speakers' production.
  – Larger breaks in the discourse (e.g. topic boundaries) show:
    • longer pauses, higher onset pitch after, final lowering/lengthening before (Lehiste, Yule, Auran, Wichmann)
  – Hierarchically higher discourse segments tend to have:
    • higher pitch range, higher mean pitch and intensity and longer preceding pauses (Hirschberg & Grosz, Tyler under revision)
Prosody in Discourse Perception

• Most use meta-linguistic judgments (except see Mayer et al. 2006, Silverman 1987)
  – Naturalness
  – Location in a paragraph

• UNC Study: Production & Perception
  • Participants saw ambiguous discourse and tried to read it to convey intended meaning.
  • Listeners were not able to retrieve speaker’s intended meaning, except for one speaker (75% success)
The Experiment: Can prosody bias interpretation?

- Prosody: two conditions
  - Coord prosody $\rightarrow$ Coord interpretation?
  - Subord prosody $\rightarrow$ Subord interpretation?

- Record neutral productions
- Synthesize contrasts in Praat
Prosody: Start with same production

• All sentences read individually in a carrier context (order randomized):

  I am going to read a sentence.  I visited my uncle in Detroit.  I just read a sentence.
  I am going to read a sentence.  I read about housing prices.  I just read a sentence.
  I am going to read a sentence.  I sat in on a history class.  I just read a sentence.
  I am going to read a sentence.  I went for a run.  I just read a sentence.
  I am going to read a sentence.  And I watched a cool documentary.  I just read a sentence.

• Intensity normed to the speaker’s overall mean (57.2 dB).
Prosody: Manipulate into Two Conditions

<table>
<thead>
<tr>
<th>S1</th>
<th>P1</th>
<th>S2</th>
<th>P2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>I sat in on a history class.</em></td>
<td><em>I read about housing prices.</em></td>
<td><em>And I watched a cool documentary.</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**COORD**

- Rise x 1.6

- [*...class*]

**SUBORD**

- Fall x 0.75

- [*...class*]

1: Pitch contour at end of Sentence 1 (S1)
Prosody: Manipulate into Two Conditions

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>P1</th>
<th>S2</th>
<th>P2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>I sat in on a history class.</em></td>
<td><em>I read about housing prices.</em></td>
<td><em>And I watched a cool documentary.</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COORD</td>
<td>Rise x 1.6</td>
<td>Rise x 1.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[...class]</td>
<td>[...pri - ces]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUBORD</td>
<td>Fall x 0.75</td>
<td>Rise x 1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[...class]</td>
<td>[...pri - ces]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2: Pitch contour at end of Sentence 2 (S2)
Prosody: Manipulate into Two Conditions

<table>
<thead>
<tr>
<th>S1</th>
<th>P1</th>
<th>S2</th>
<th>P2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>I sat in on a history class.</em></td>
<td><em>I read about housing prices.</em></td>
<td><em>And I watched a cool documentary.</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**COORD**
- Rise x 1.6
- Rise x 1.3
- 920ms
- [...class]
- [...pri - ces]

**SUBORD**
- Fall x 0.75
- Rise x 1.1
- 400ms
- [...class]
- [...pri - ces]

3: Pause 1, duration of silence between S1 and S2
Prosody: Manipulate into Two Conditions

<table>
<thead>
<tr>
<th>S1</th>
<th>P1</th>
<th>S2</th>
<th>P2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>I sat in on a history class.</td>
<td></td>
<td>I read about housing prices.</td>
<td></td>
<td>And I watched a cool documentary.</td>
</tr>
</tbody>
</table>

**COORD**
- Rise x 1.6
- 920ms
- [...class]

**SUBORD**
- Fall x 0.75
- 400ms
- [...class]

- Rise x 1.3
- 400ms
- [...prices]

- Rise x 1.1
- 20ms
- [...prices]

4: Pause 2, duration of silence between S2 and S3
### Prosody: Manipulate into Two Conditions

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>P1</th>
<th>S2</th>
<th>P2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I sat in on a history class.</td>
<td>I read about housing prices.</td>
<td>And I watched a cool documentary.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COORD</td>
<td>Rise x 1.6</td>
<td>Rise x 1.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[...class]</td>
<td>[...prices]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUBORD</td>
<td>Fall x 0.75</td>
<td>Rise x 1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[...class]</td>
<td>[...prices]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5: Mean pitch and intensity on S2 and S3

Mean pitch & intensity x 1.1 for S2 and S3

Mean pitch & intensity x 0.9 for S2 and S3
The Experiment: Can prosody bias interpretation?

- 40 participants (University of Michigan Psychology subject pool)
- 48 discourses (102 discourses normed, and 48 most ambiguous ones used)
Procedure

**Interpretation Question**
Did Sally mean that she read about housing prices and watched a cool documentary in the history class?

![Discourse](image)

**Confidence Question**
How confident are you in that choice?

**Comprehension Question**
Did Sally pick up some beer?
Can prosody bias discourse interpretation?

• Coord prosody $\rightarrow$ Coord interpretation?
• Subord prosody $\rightarrow$ Subord interpretation?

• Statistical model
  – Generalized Linear Mixed Model for binary judgment data
  – Linear Mixed Model for confidence data
Results: Interpretation

Coord interpretation

Subord interpretation
Results: Confidence

[Bar chart showing mean confidence levels for mismatch and match conditions with error bars indicating 95% CI.]
Is this result replicable with subjects from Mechanical Turk?

• Amazon Mechanical Turk is an online labor marketplace; you post tasks, and participants (“turkers”) complete them for a set fee.

• The exact same survey was administered to participants from Mechanical Turk.
Results: Psych Pool & Mech Turk

Coord interpretation

Subord interpretation
Which prosodic contrasts are having the effect on interpretation?

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>P1</th>
<th>S2</th>
<th>P2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>I sat in on a history class.</em></td>
<td><em>I read about housing prices.</em></td>
<td><em>And I watched a cool documentary.</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 1: Pitch contour at end of Sentence 1 (S1)
- 2: Pitch contour at end of Sentence 2 (S2)
- 3: Pause 1, duration of silence between S1 and S2
- 4: Pause 2, duration of silence between S2 and S3
- 5: Mean pitch and intensity on S2 and S3
Results: Combinations of prosodic contrasts

1: Pitch contour at end of Sentence 1 (S1)
2: Pitch contour at end of Sentence 2 (S2)
3: Pause 1, duration of silence between S1 and S2
4: Pause 2, duration of silence between S2 and S3
5: Mean pitch and intensity on S2 and S3
Results: Combinations of prosodic contrasts

1: Pitch contour at end of Sentence 1 (S1)
2: Pitch contour at end of Sentence 2 (S2)
3: Pause 1, duration of silence between S1 and S2
4: Pause 2, duration of silence between S2 and S3
5: Mean pitch and intensity on S2 and S3
Results: Combinations of prosodic contrasts (All)

WITH S1 terminal pitch

- 1: Pitch contour at end of Sentence 1 (S1)
- 2: Pitch contour at end of Sentence 2 (S2)
- 3: Pause 1, duration of silence between S1 and S2
- 4: Pause 2, duration of silence between S2 and S3
- 5: Mean pitch and intensity on S2 and S3

WITHOUT S1 terminal pitch

Error bars: 95% CI
I sat in on a history class. I read about housing prices. And I watched a cool documentary.
Conclusions

• Prosody can bias the interpretation of ambiguous discourse.
  – Lexical and syntactic information held constant, so contrast is in relationship between sentences.
  – Listeners are more confident when they choose the interpretation predicted by the prosody.

• The overall effect can be replicated with Mechanical Turk participants, though it is a weaker effect.

• The prosodic manipulation driving the effect is the terminal pitch on sentence 1; the rest seem irrelevant.
Thank you!

• Comments and suggestions welcome: jctyler@umich.edu