

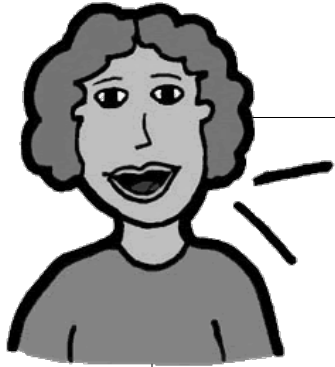
# Prosody and Listeners' Interpretation of Ambiguous Discourse

Joseph Tyler

University of Michigan

[jctyler@umich.edu](mailto:jctyler@umich.edu)

**Sally**



***I sat in on a history class.  
I read about housing prices.  
And I watched a cool documentary.***

Did Sally mean that she read about housing prices  
and watched a cool documentary...

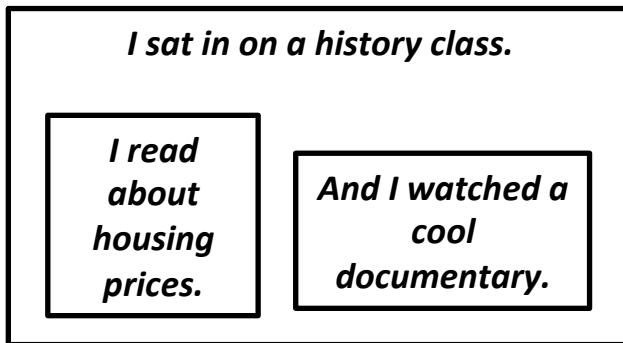
in history class?

separate from history class?

# Two Interpretations

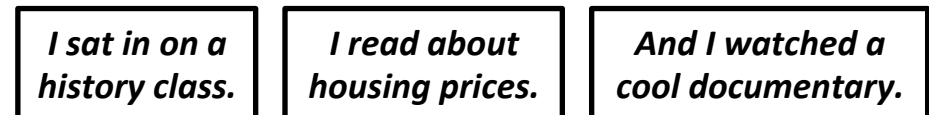
***I sat in on a history class. [S1]***  
***I read about housing prices. [S2]***  
***And I watched a cool documentary. [S3]***

## SUBORD



S1 event elaborated by S2 and S3

## COORD



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 → Coord interpretation?
  - Subord prosody → 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>I am going to read a sentence.</i>	<b><i>I visited my uncle in Detroit.</i></b>	<i>I just read a sentence.</i>
<i>I am going to read a sentence.</i>	<b><i>I read about housing prices.</i></b>	<i>I just read a sentence.</i>
<i>I am going to read a sentence.</i>	<b><i>I sat in on a history class.</i></b>	<i>I just read a sentence.</i>
<i>I am going to read a sentence.</i>	<b><i>I went for a run.</i></b>	<i>I just read a sentence.</i>
<i>I am going to read a sentence.</i>	<b><i>And I watched a cool documentary.</i></b>	<i>I just read a sentence.</i>

...

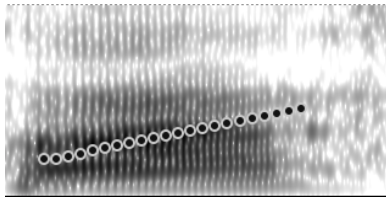
- Intensity normed to the speaker's overall mean (57.2 dB).

# Prosody: Manipulate into Two Conditions

S1	P1	S2	P2	S3
<i>I sat in on a history class.</i>		<i>I read about housing prices.</i>		<i>And I watched a cool documentary.</i>

COORD

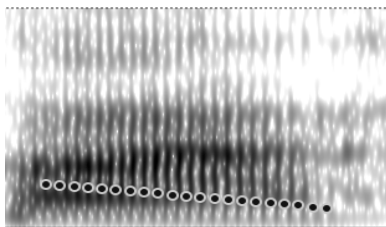
Rise x 1.6



[...class]

SUBORD

Fall x 0.75



[...class]

1: Pitch contour at end of Sentence 1 (S1)

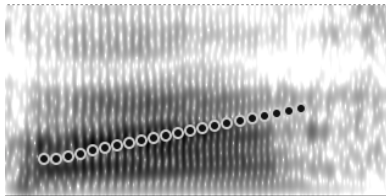


# Prosody: Manipulate into Two Conditions

S1	P1	S2	P2	S3
<i>I sat in on a history class.</i>		<i>I read about housing prices.</i>		<i>And I watched a cool documentary.</i>

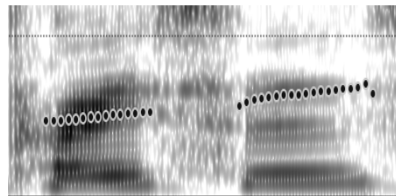
COORD

Rise x 1.6



[...class]

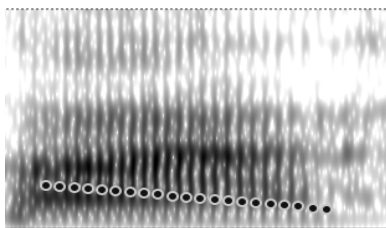
Rise x 1.3



[...pri - ces]

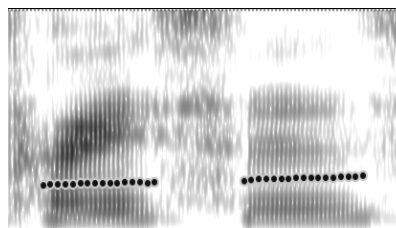
SUBORD

Fall x 0.75



[...class]

Rise x 1.1



[...pri - ces]

2: Pitch contour at end of Sentence 2 (S2)

# Prosody: Manipulate into Two Conditions

S1	P1	S2	P2	S3
----	----	----	----	----

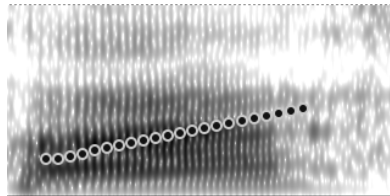
*I sat in on a  
history class.*

*I read about  
housing prices.*

*And I watched a  
cool documentary.*

COORD

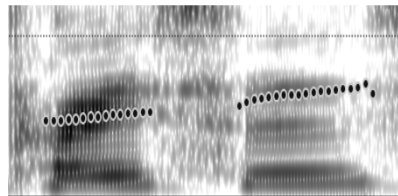
Rise x 1.6



920ms

[...class]

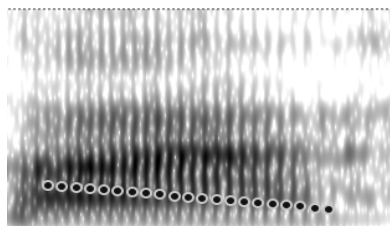
Rise x 1.3



[...pri - ces]

SUBORD

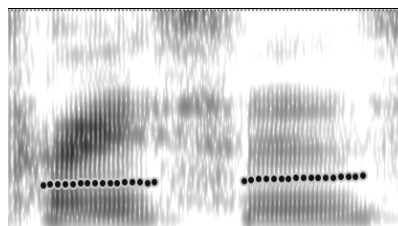
Fall x 0.75



400ms

[...class]

Rise x 1.1



[...pri - ces]

3: Pause 1, duration of silence between S1 and S2

# Prosody: Manipulate into Two Conditions

S1	P1	S2	P2	S3
----	----	----	----	----

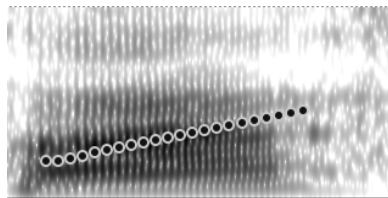
*I sat in on a  
history class.*

*I read about  
housing prices.*

*And I watched a  
cool documentary.*

COORD

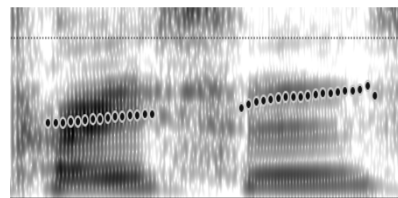
Rise x 1.6



920ms

[...class]

Rise x 1.3

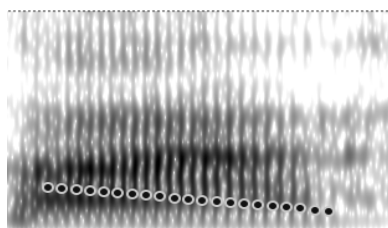


400ms

[...pri - ces]

SUBORD

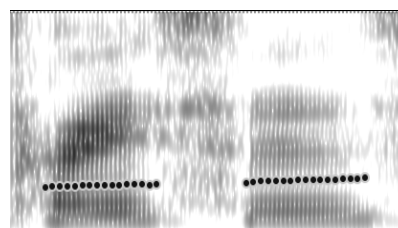
Fall x 0.75



400ms

[...class]

Rise x 1.1



20ms

[...pri - ces]

4: Pause 2, duration of silence between S2 and S3

# Prosody: Manipulate into Two Conditions

S1	P1	S2	P2	S3
----	----	----	----	----

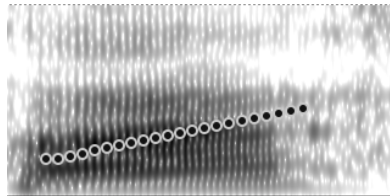
*I sat in on a history class.*

*I read about housing prices.*

*And I watched a cool documentary.*

COORD

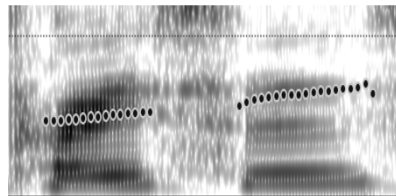
Rise x 1.6



920ms

[...class]

Rise x 1.3



400ms

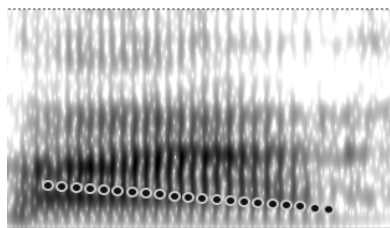
[...pri - ces]



Mean pitch & intensity x 1.1  
for S2 and S3

SUBORD

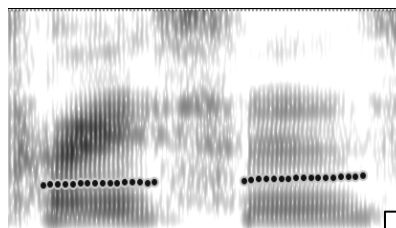
Fall x 0.75



400ms

[...class]

Rise x 1.1



20ms

[...pri - ces]



Mean pitch & intensity x 0.9  
for S2 and S3

5: Mean pitch and intensity on 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



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

Yes

No

## Confidence Question

How confident are you in that choice?

## Comprehension Question

Did Sally pick up some beer?

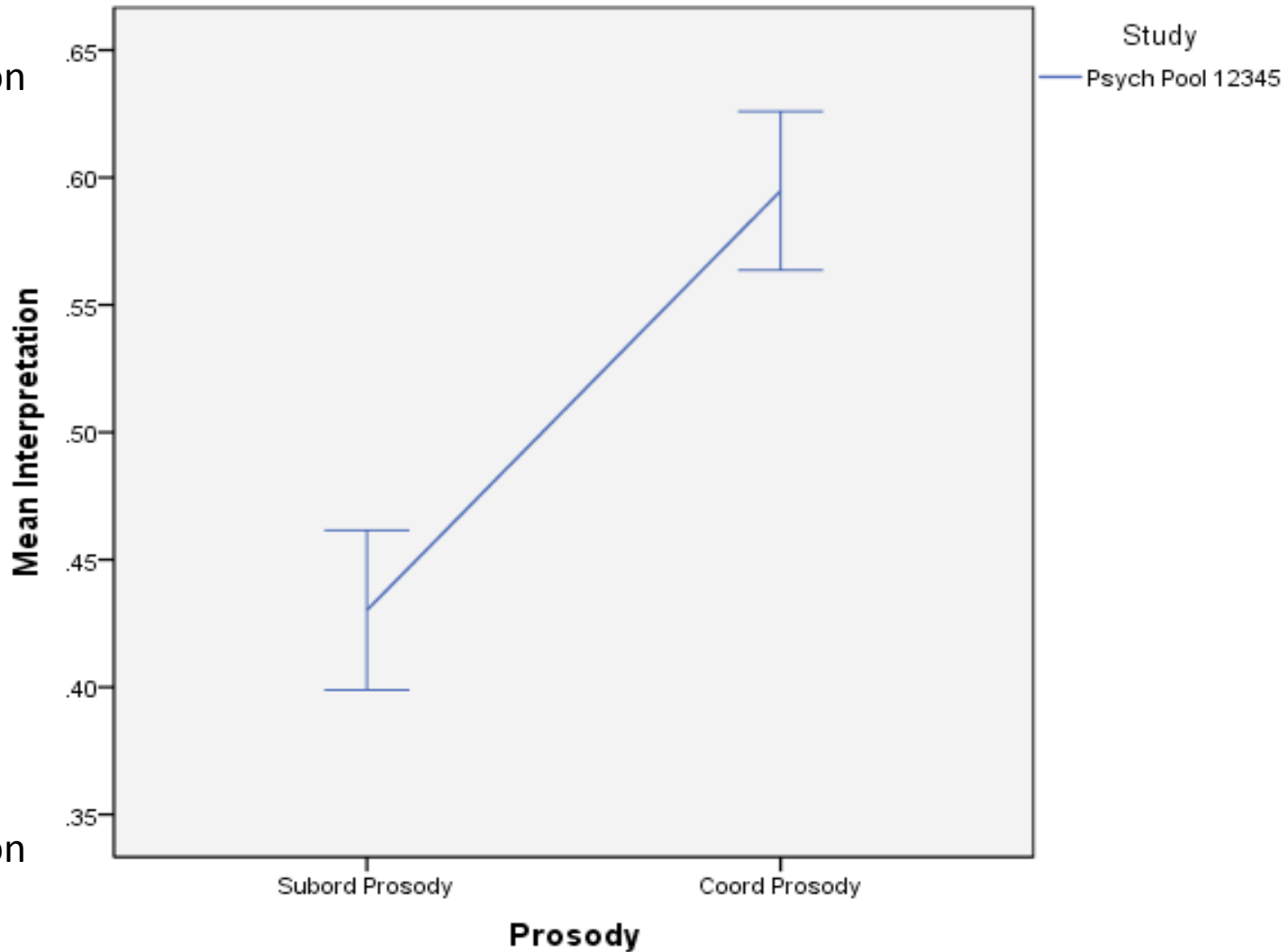
# Can prosody bias discourse interpretation?

- Coord prosody → Coord interpretation?
- Subord prosody → Subord interpretation?
- Statistical model
  - Generalized Linear Mixed Model for binary judgment data
  - Linear Mixed Model for confidence data

# Results: Interpretation

Coord  
interpretation

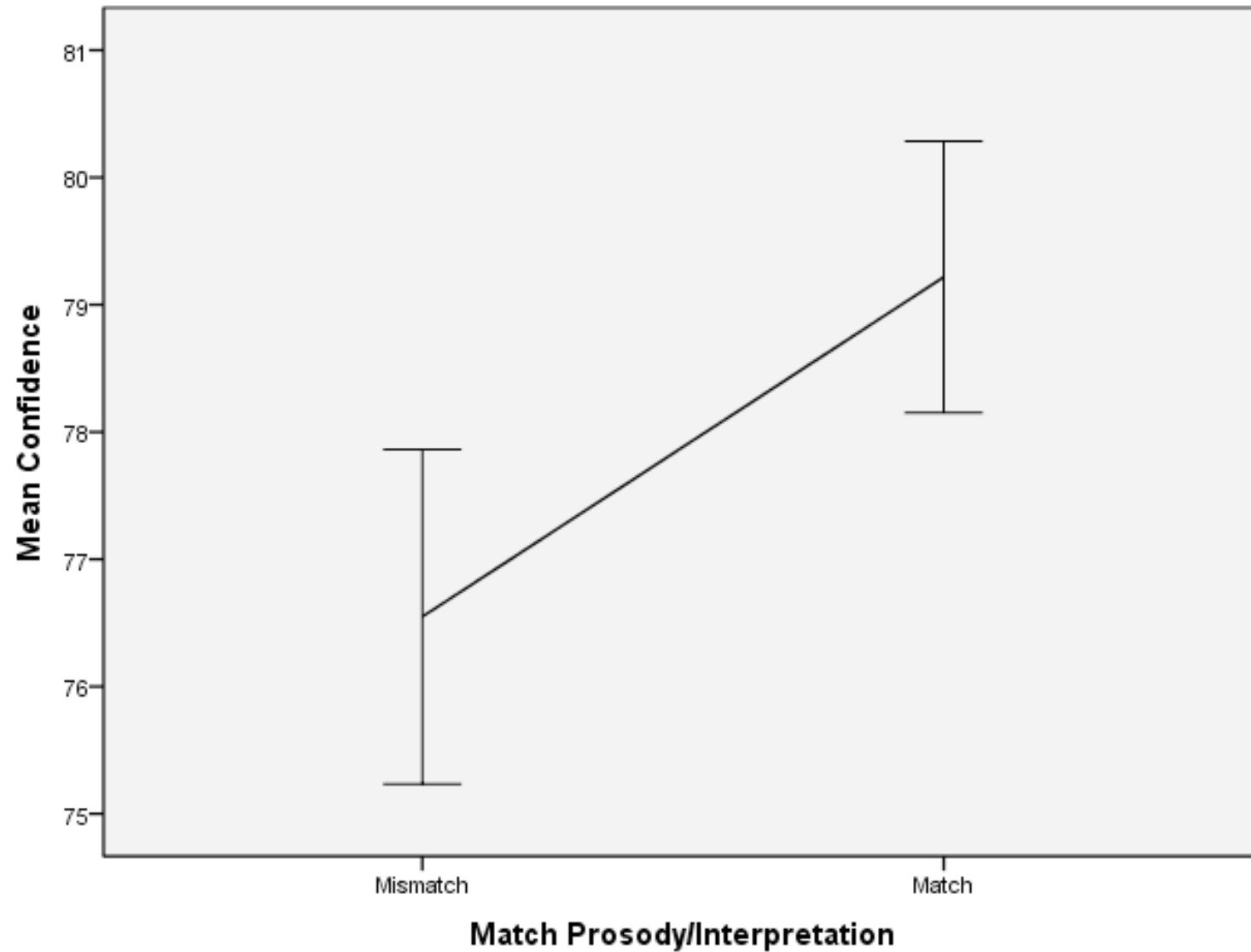
Subord  
interpretation



Error bars: 95% CI



# Results: Confidence



Error bars: 95% CI

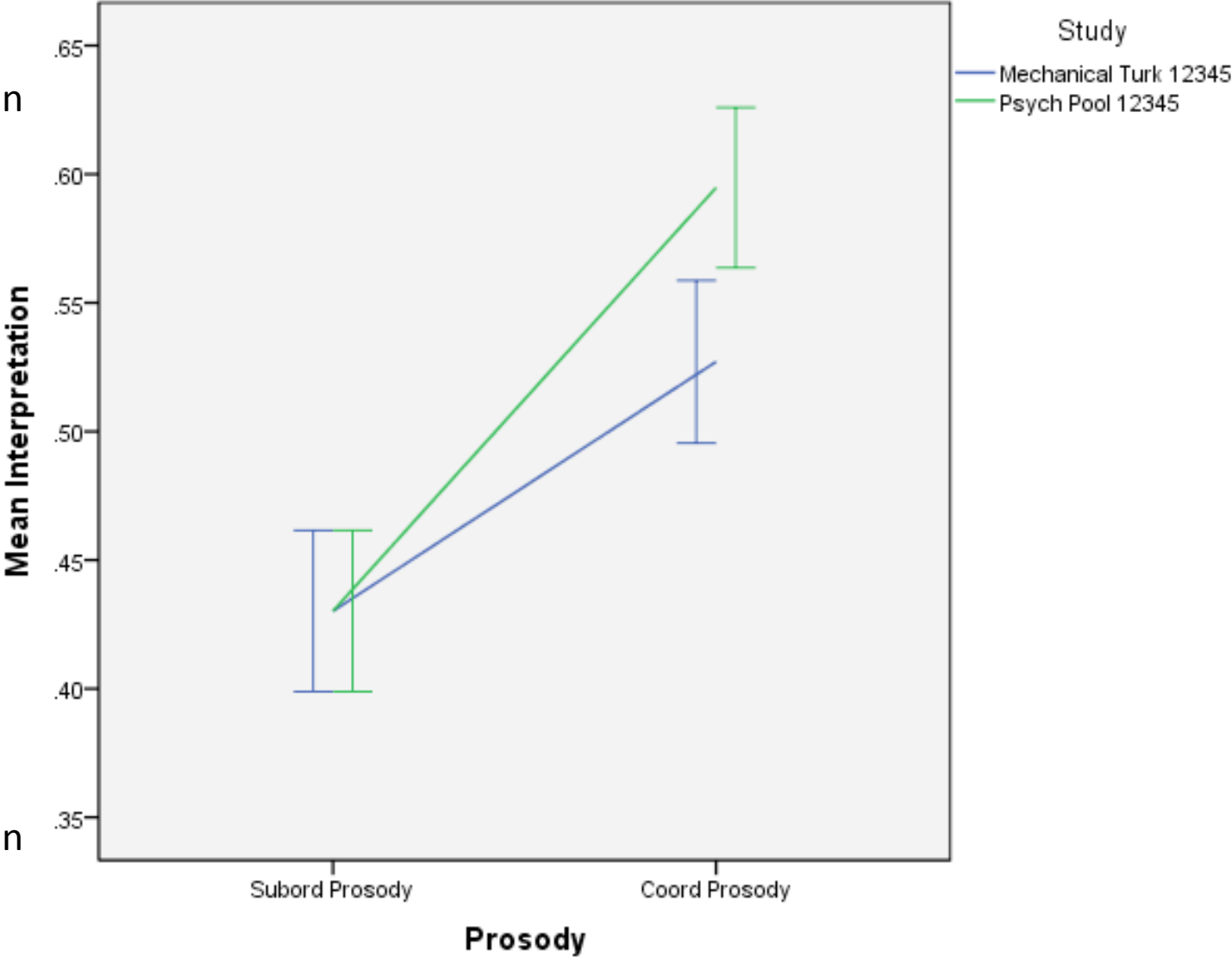
# Is this result replicable with subjects from Mechanical Turk?



- 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

Error bars: 95% CI

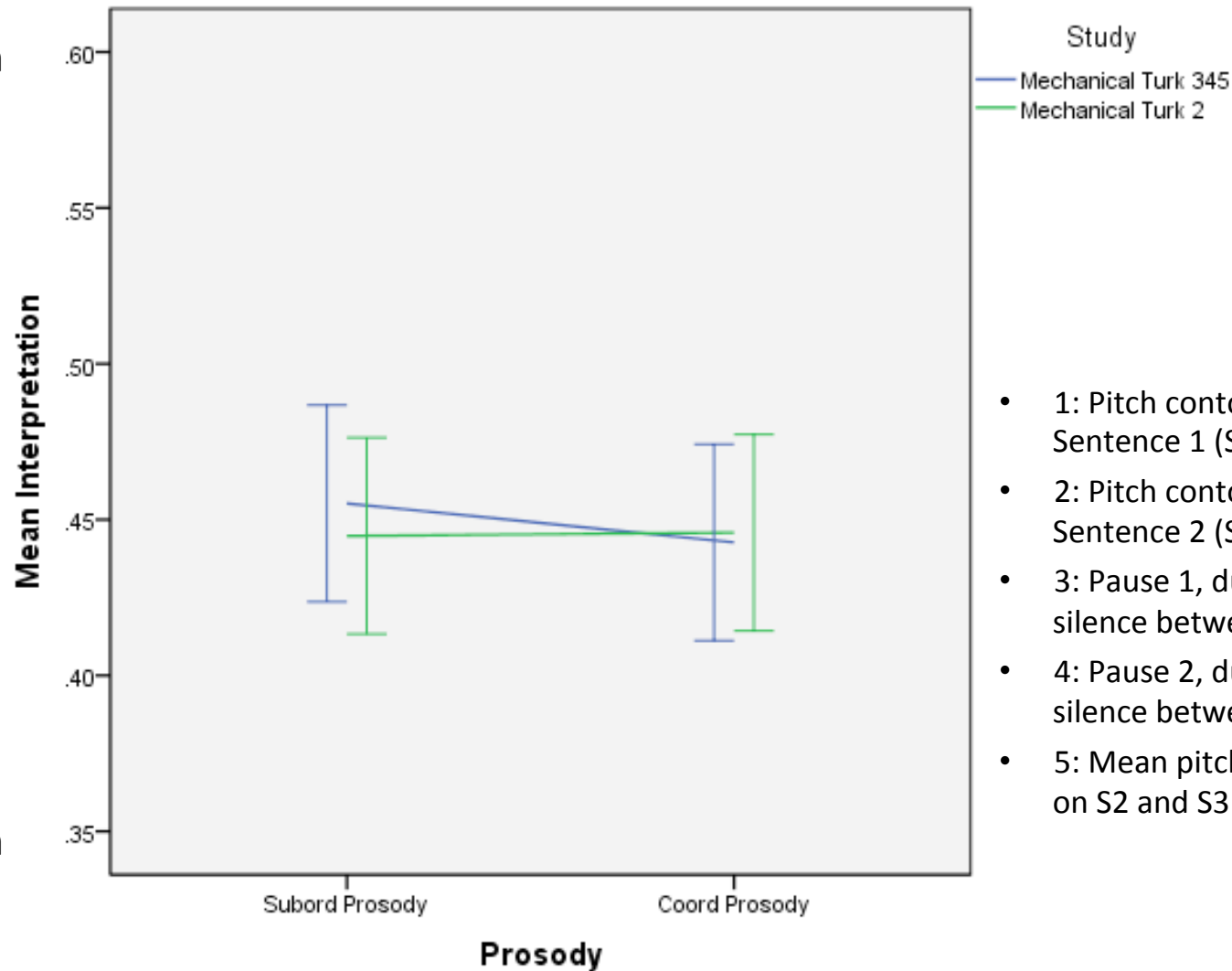
# Which prosodic contrasts are having the effect on interpretation?

S1	P1	S2	P2	S3
<i>I sat in on a history class.</i>		<i>I read about housing prices.</i>		<i>And I watched a cool documentary.</i>

- 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

Coord  
interpretation



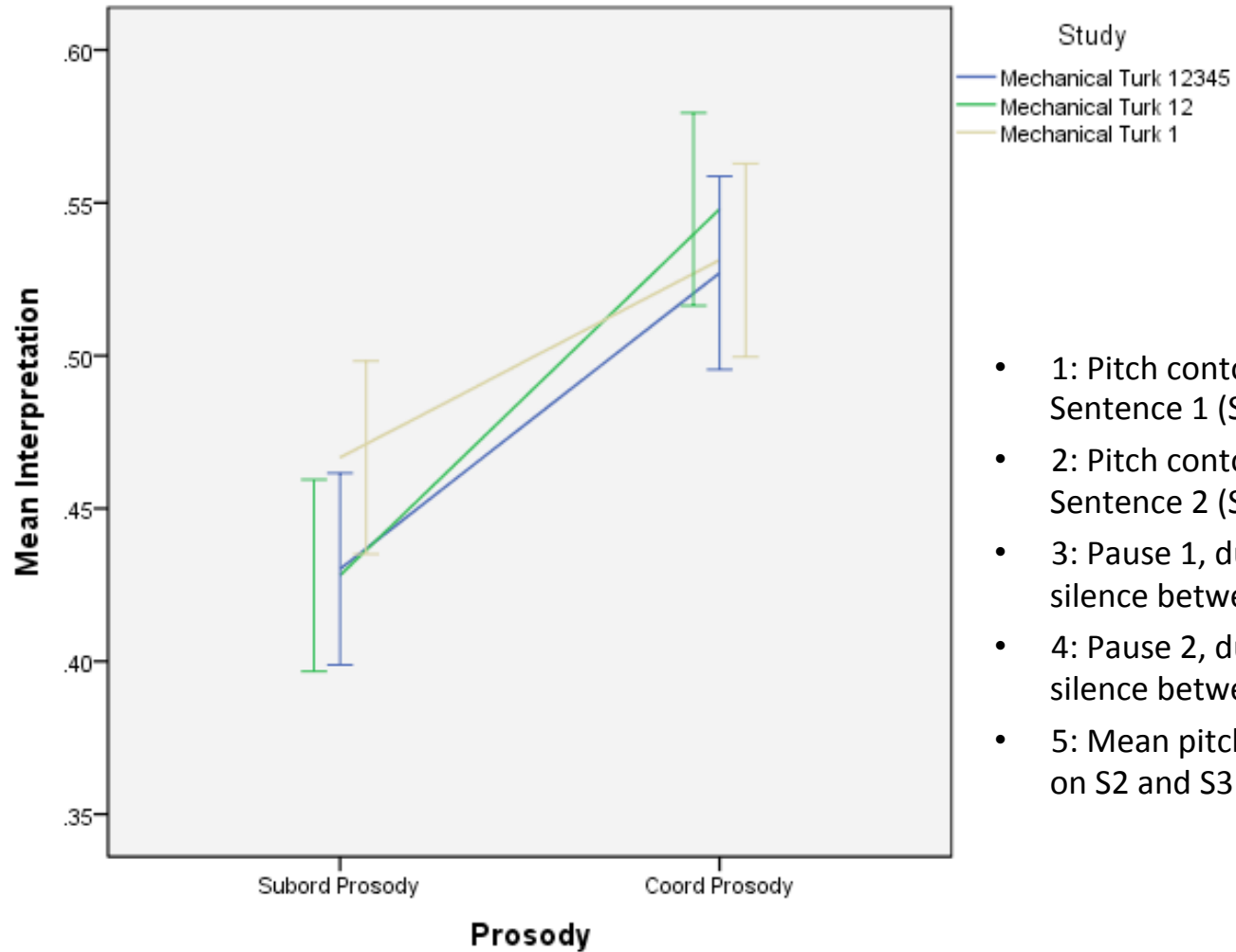
Subord  
interpretation

- 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

Error bars: 95% CI

# Results: Combinations of prosodic contrasts

Coord  
interpretation



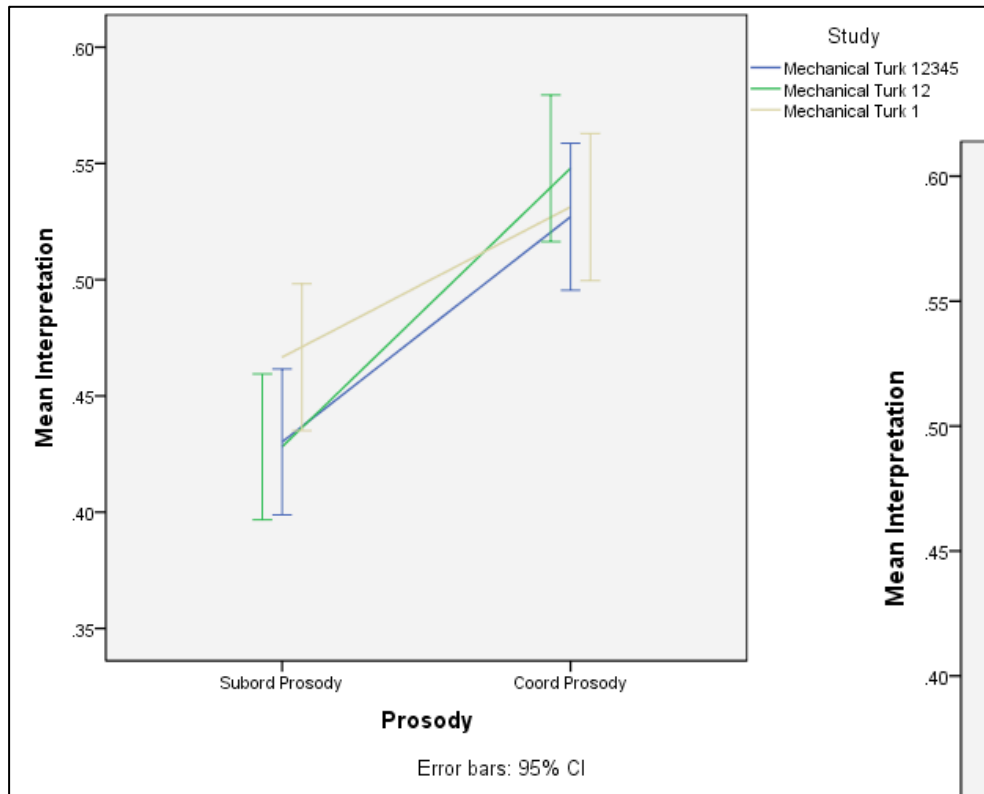
Subord  
interpretation

- 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

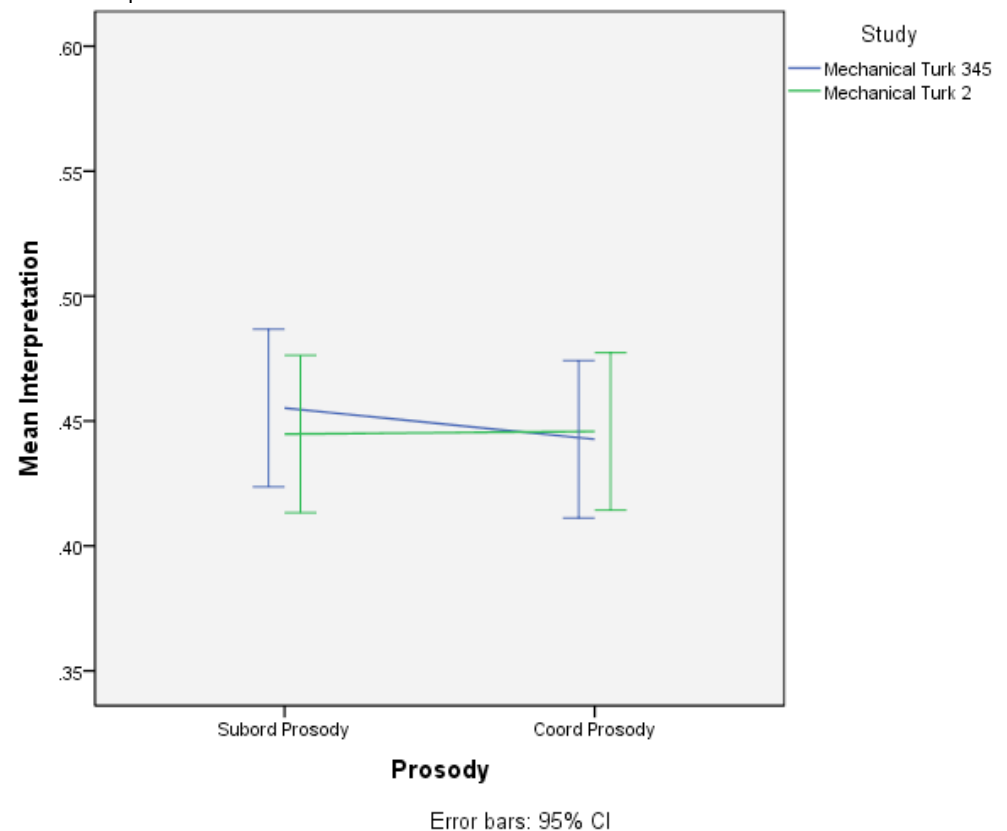
Error bars: 95% CI

# Results: Combinations of prosodic contrasts (All)

WITH S1 terminal pitch

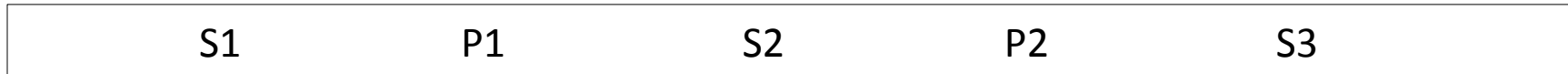


WITHOUT S1 terminal pitch



- 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

# S1 Terminal Pitch Drives Effect



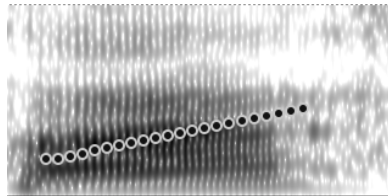
*I sat in on a history class.*

*I read about housing prices.*

*And I watched a cool documentary.*

COORD

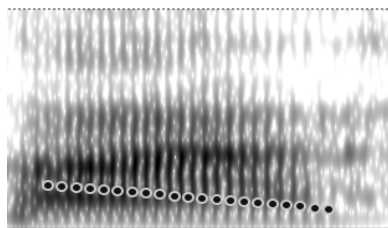
Rise x 1.6



[...class]

SUBORD

Fall x 0.75

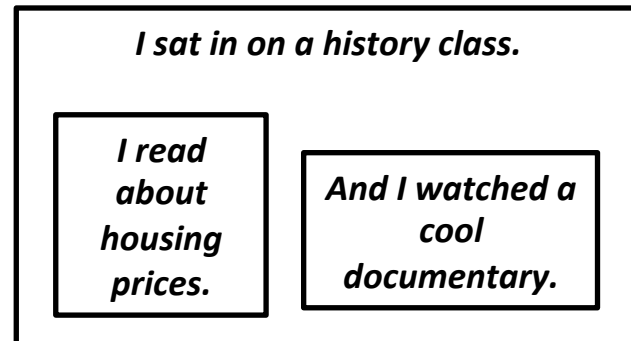


[...class]

*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](mailto:jctyler@umich.edu)

