Choosing among alternatives:

Conjunction variability comes from both inference and the semantics of discourse adverbials

Hannah Rohde, Alexander Johnson,
Nathan Schneider, & Bonnie Webber

Discourse coherence

Recipe for whipped cream frosting:
    Put cream cheese and whipping cream into a bowl.
    (then)
    Add sugar and vanilla.
    (then)
    Beat the mixture until the cream can hold a stiff peak.
    (then)
    Cover cakes with this frosting that won't melt at room temperature.

    Otherwise **you'll be left with soggy cupcakes.**

→ Some relations can be left implicit; others can't.

(Asher & Lascarides, 2003; Hobbs, 1979; Kehler, 2002; Mann &
Thompson, 1988; Prasad et al, 2014; Roberts, 1996; Sanders et al., 1992)

This talk: Recovering implicit relations

▸ A puzzle for existing models of coherence relations

▸ Applications of coherence inferences

▸ Conjunction-insertion experiment

→ Results show role for inference alongside explicit cues

A puzzle

▸ Deduction of implicit information from juxtaposed sentences

It's too far to walk. Let's take the bus.

Infer alternatives: walk/bus as means of transport
Infer causal relation: too far, therefore bus
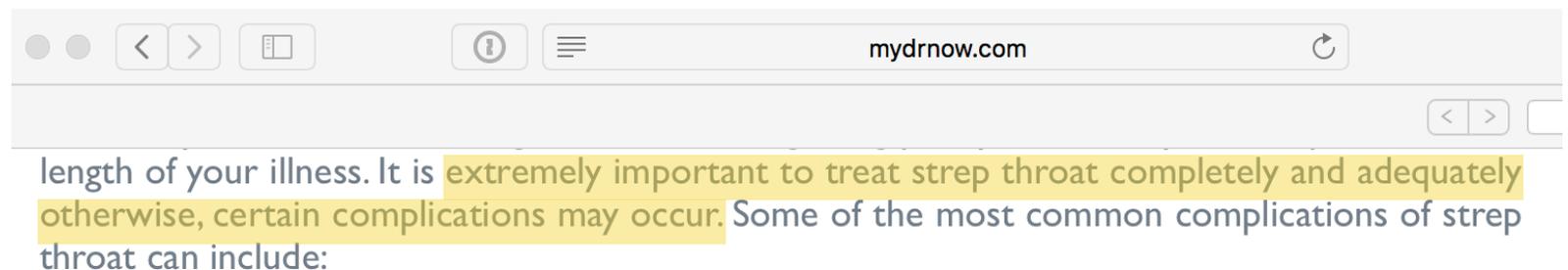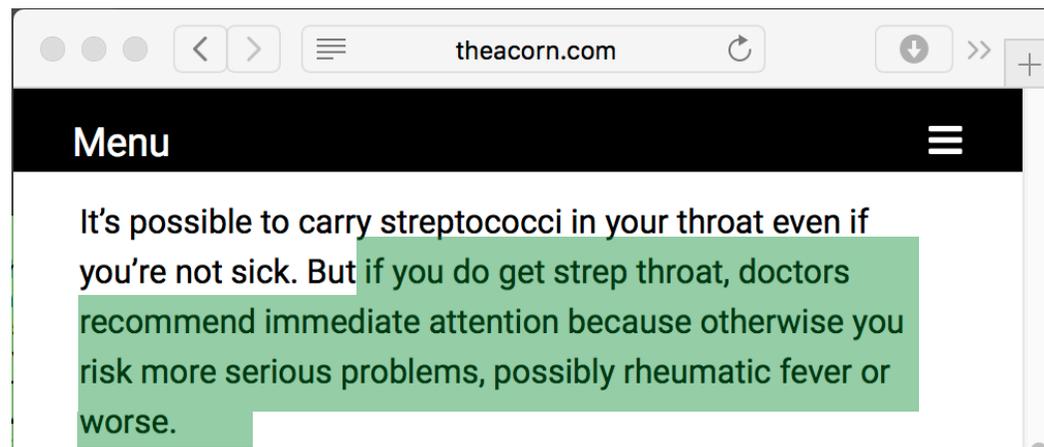
It's too far to walk so let's take the bus.

▸ **Assumption:** A passage marks its coherence relation either explicitly or implicitly — i.e., if explicit connective is present, no need for further inference about additional relations.

It's too far to walk. so? Instead let's take the bus.

Coherence relations in NLP
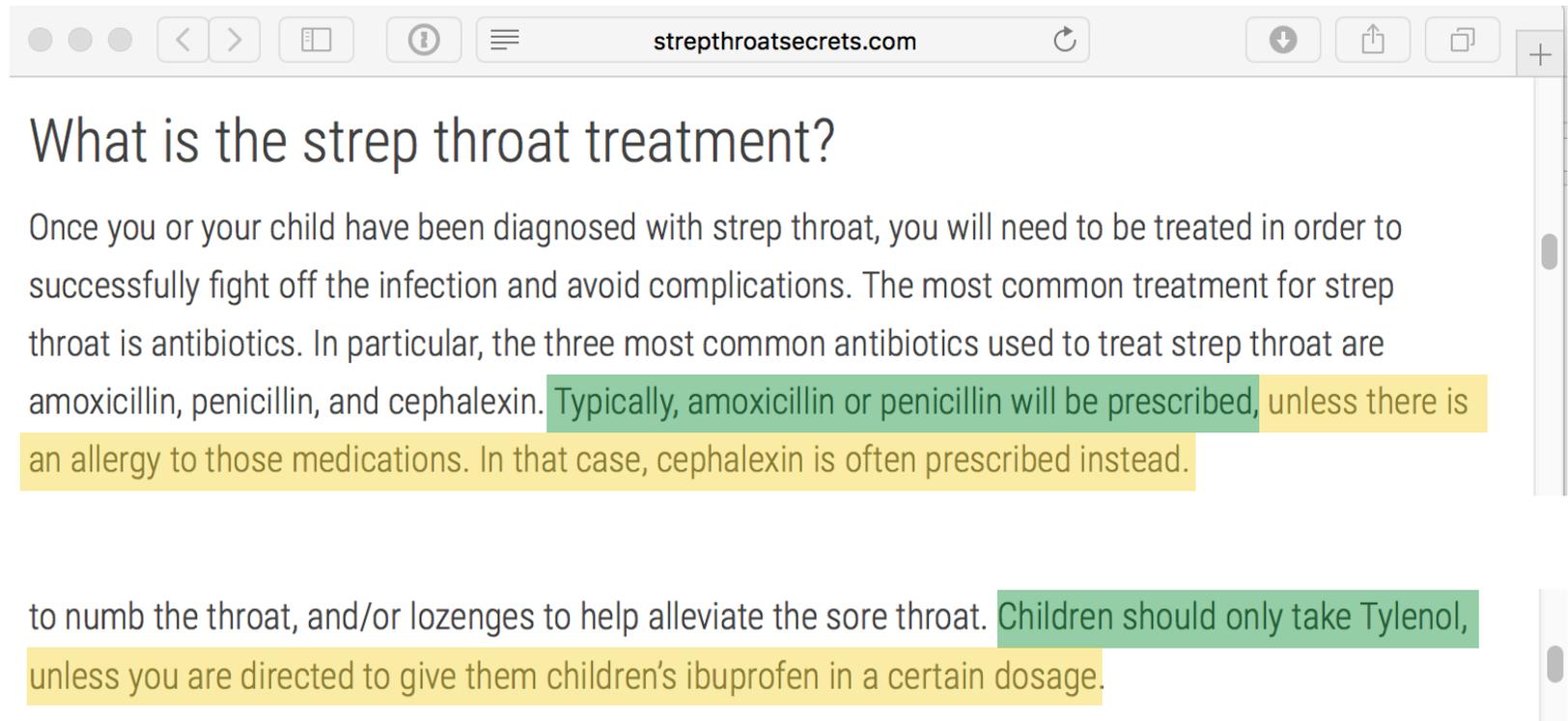
▶ Question-answering

Query: "why treat strep throat?"

It's possible to carry streptococci in your throat even if you're not sick. But if you do get strep throat, doctors recommend immediate attention because otherwise you risk more serious problems, possibly rheumatic fever or worse.

length of your illness. It is extremely important to treat strep throat completely and adequately otherwise, certain complications may occur. Some of the most common complications of strep throat can include:

Coherence relations in text

▶ Question-answering

> **Query: "why treat strep throat?"**

> **Query: "how to treat strep throat?"**

strepthroatsecrets.com

What is the strep throat treatment?

Once you or your child have been diagnosed with strep throat, you will need to be treated in order to successfully fight off the infection and avoid complications. The most common treatment for strep throat is antibiotics. In particular, the three most common antibiotics used to treat strep throat are amoxicillin, penicillin, and cephalexin. Typically, amoxicillin or penicillin will be prescribed, unless there is an allergy to those medications. In that case, cephalexin is often prescribed instead.

to numb the throat, and/or lozenges to help alleviate the sore throat. Children should only take Tylenol, unless you are directed to give them children's ibuprofen in a certain dosage.

Coherence relations in text

▸ Question-answering

> **Query:** **"why treat strep throat?"**

> **Query:** **"how to treat strep throat?"**

→ Extraction of best answer may depend on linked clauses

→ Links may not always be explicit

Coherence relations in text

▶ Question-answering

▶ Text generation, automatic summarisation:
    Systems must decide what to make explicit to sound natural.

▶ Coreference resolution
    Best antecedent may vary across coherence relations.

▶ Given this utility, large-scale annotated resources have been developed.

Back to the puzzle

▸ **Assumption:** A passage marks its coherence relation either explicitly or implicitly — i.e., if explicit connective is present, no need for further inference about additional relations.

> so?
> It's too far to walk. ∧Instead let's take the bus.

▸ **Question:** Are additional inferences necessary even when an explicit cue is present?

Psycholinguistic studies

1. **Do inferable discourse relations hold when a discourse adverbial is already present?**
   - → Yes, adverbials license co-occurring conjunctions
2. **How to characterise discourse adverbials with respect to inferred relations?**
   - → Not predictable from adverbial or semantic class
   - → More than one valid connection in some cases
3. **How to account for unexpected combinations?**
   - → Multiple simultaneous sources of coherence

Conjunction-insertion (presented TextLink 2015)

→ Update: current dataset of judgments for 50 adverbials, each in 50+ passages, each passage judged by 28 people... 70,000+ data points

(Rohde et al., 2015, 2016, 2017)

Passages in dataset

▶ Materials: for each adverbial, 50+ passages (mostly) from NYTimes Annotated Corpus (Sandhaus, 2008)

▶ Half originally explicit

> **"Nervous? No, my leg's not shaking," said Griffey, who caused everyone to laugh // _____ indeed his right foot was shaking.**

Author=BECAUSE

▶ Half originally implicit

> **Sellers are usually happy, too // _____ after all they are the ones leaving with money.**

Author=NONE

Adverbials include: ACTUALLY, AFTER ALL, FIRST OF ALL, FOR EXAMPLE, FOR INSTANCE, IN FACT, IN OTHER WORDS, INDEED, INSTEAD, NEVERTHELESS, NONETHELESS, ON THE ONE HAND, ON THE OTHER HAND, OTHERWISE, SPECIFICALLY, THEN, THEREFORE, THUS, …

Judgments from naive annotators

▶ Each passage viewed by 28 participants.

▶ Instructions:
Find conjunction to 'best reflect meaning of connection' between text spans

ConnText   University of Edinburgh

Trial Show Instructions

I don't mind **walking //** _____ **in fact** it's good exercise

Conjunction:
- ○ Or
- ○ But
- ○ Because
- ○ None at all
- ○ So
- ○ And
- ○ Before
- ○ Other word or phrase

Once you have made your selections, press submit to complete the trial. To share additional comments about this trial, please click here.

Submit

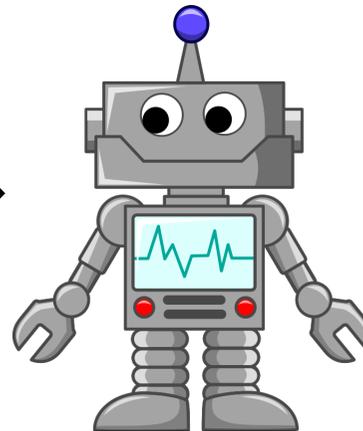▶ Catch trials

You can lead a horse to water // ___ you can't make it drink

Hypotheses for implicit passages

▶ **Variability across adverbials:** Do implicit passages pattern uniformly or vary across adverbials (by semantic type)?

▶ **Variability within adverbials:** Does the adverbial predict the same conjunction for all passages?
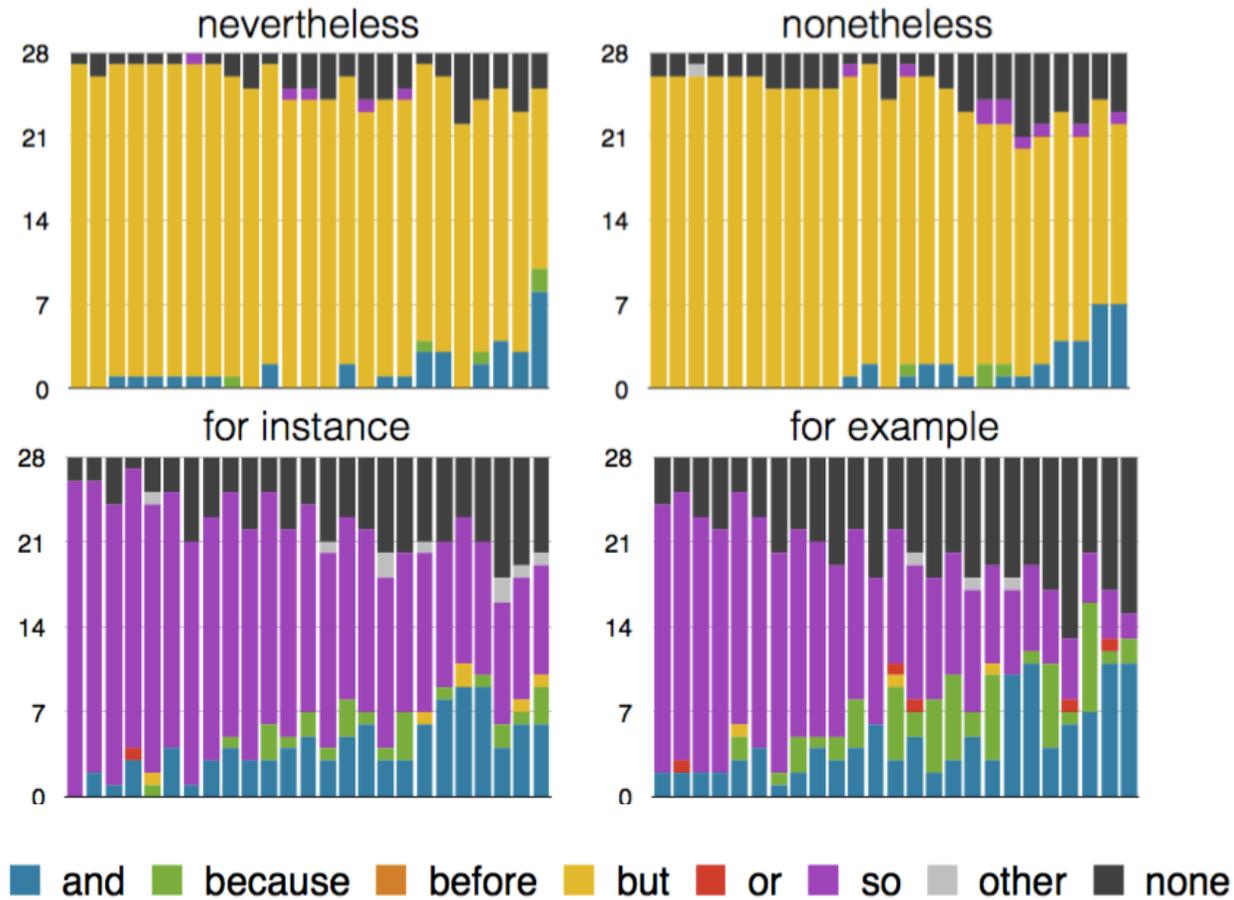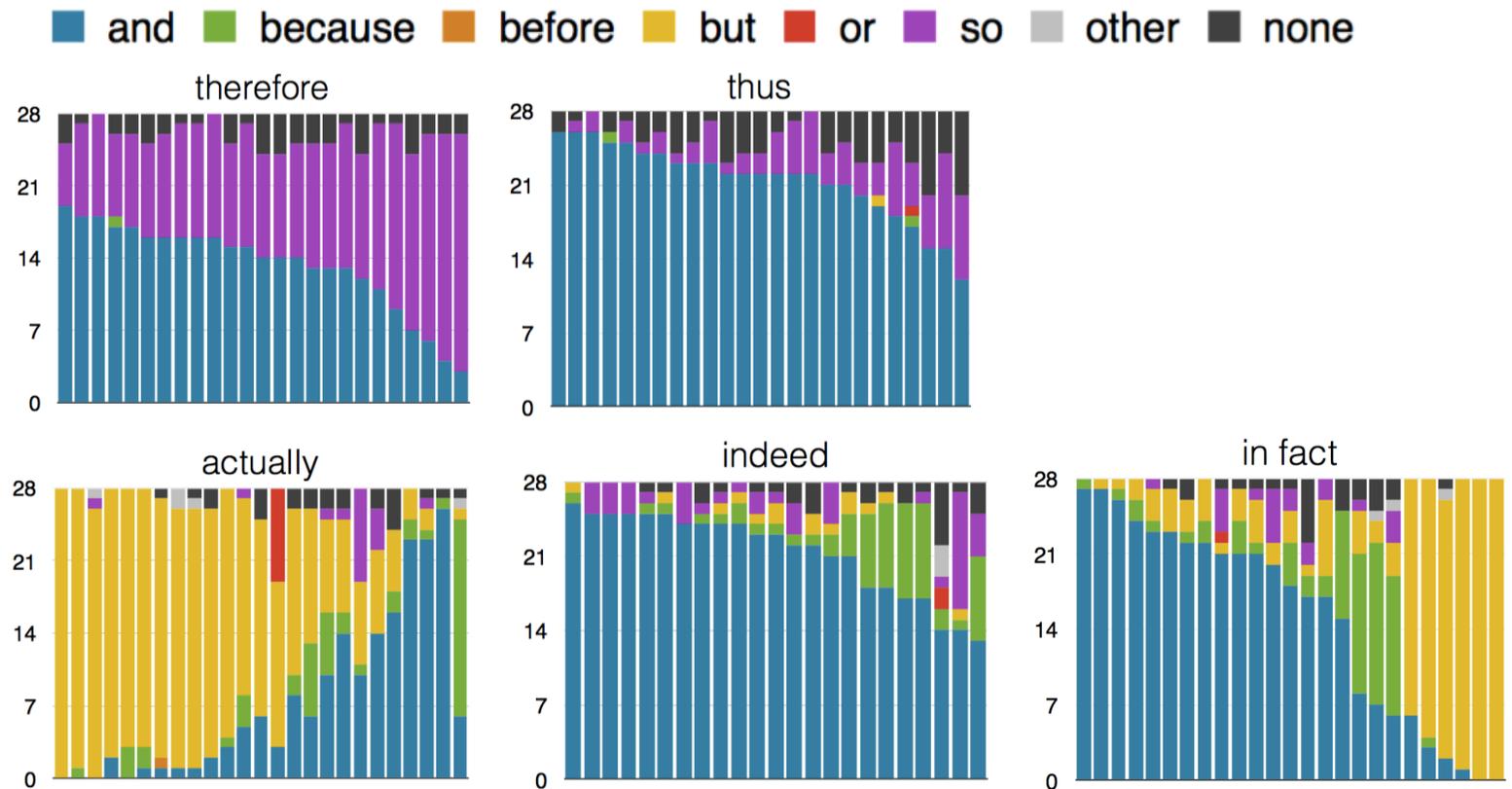
▶ If deterministic →

▶ If not →

Implicit passages

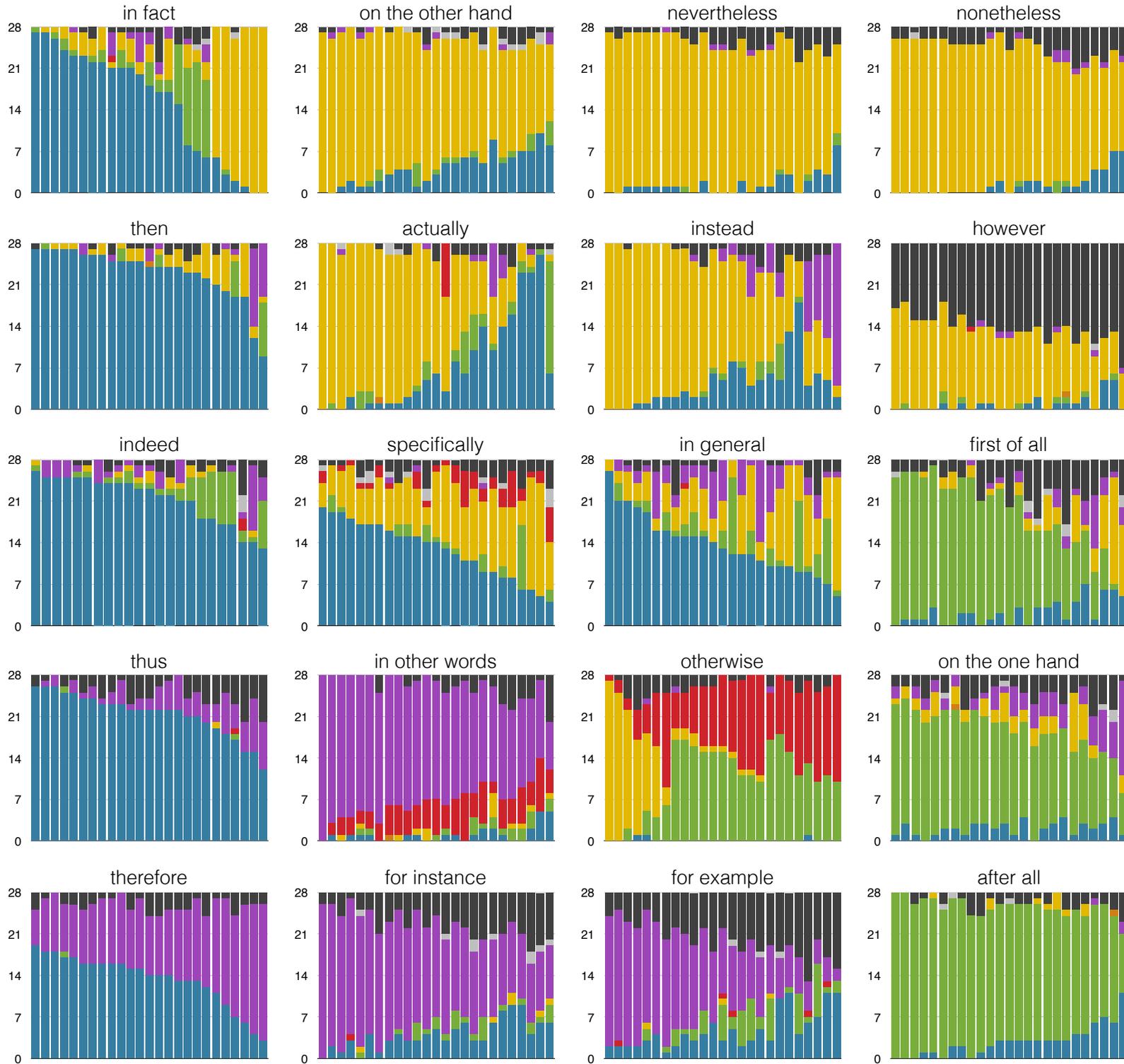▶ On one hand, we saw some consistency in semantically related adverbial pairs.

Implicit passages

▶ But also divergence for near synonyms or for adverbials of a similar type (e.g., modal stance)

▶ Adverbial itself matters, as does passage content.

Legend: and, because, before, but, or, so, other, none

in fact · on the other hand · nevertheless · nonetheless
then · actually · instead · however
indeed · specifically · in general · first of all
thus · in other words · otherwise · on the one hand
therefore · for instance · for example · after all

Cases of disagreement

▸ Adverbial-specific patterns arise: e.g., Author~Participant divergence with *otherwise*

> **"The Ravitch camp has had about 25 fund-raisers and has scheduled 20 more. Thirty others are in various stages of planning," Ms. Marcus said. "It has to be highly organized // _____ otherwise it's total chaos," she added.**
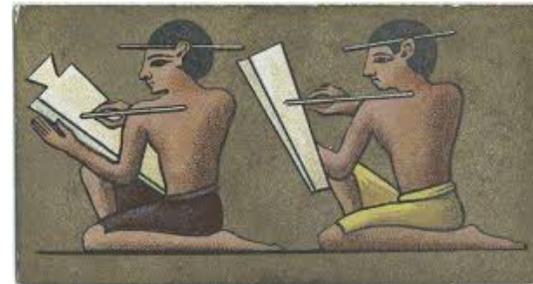
Author=OR
17 Participants=OR
11 Participants=BECAUSE

▸ Not noise

▸ Not evidence of ambiguity

▸ Improbable combinations, but perfectly fine

Implications for annotation efforts

▶ Disagreements are not errors, contra prior work on:

  ▶ Demonstrations that naive annotators can reliably infer discourse relations (Scholman et al., 2016)

  ▶ Corrections for unreliable/inattentive participants (Hovy et al. 2013, Passonneau & Carpenter 2014)

  ▶ Importance of many annotators for reducing bias (Artstein & Poesio, 2005, 2008)

▶ All with similar assumption of a single correct answer

Summary so far

▶ **Multiple connectives:** Establish necessity of entertaining implicit relations when adverbial is present

▶ **Context sensitivity:** Adverbial alone does not completely predict discourse relation

▶ **Informative disagreement:** Demonstrate possibility of divergent valid annotations

e.g., Adverbials that encode 'alternative'

▶ Lexical semantics of certain adverbials licenses OR

▶ Inference from passage content licenses BECAUSE, SO, BUT, etc.

Gouges are deep scratches that must be filled as well as colored _____ otherwise they will collect dirt and become permanently discolored.

→ *otherwise* encodes 'otherness' (OR)
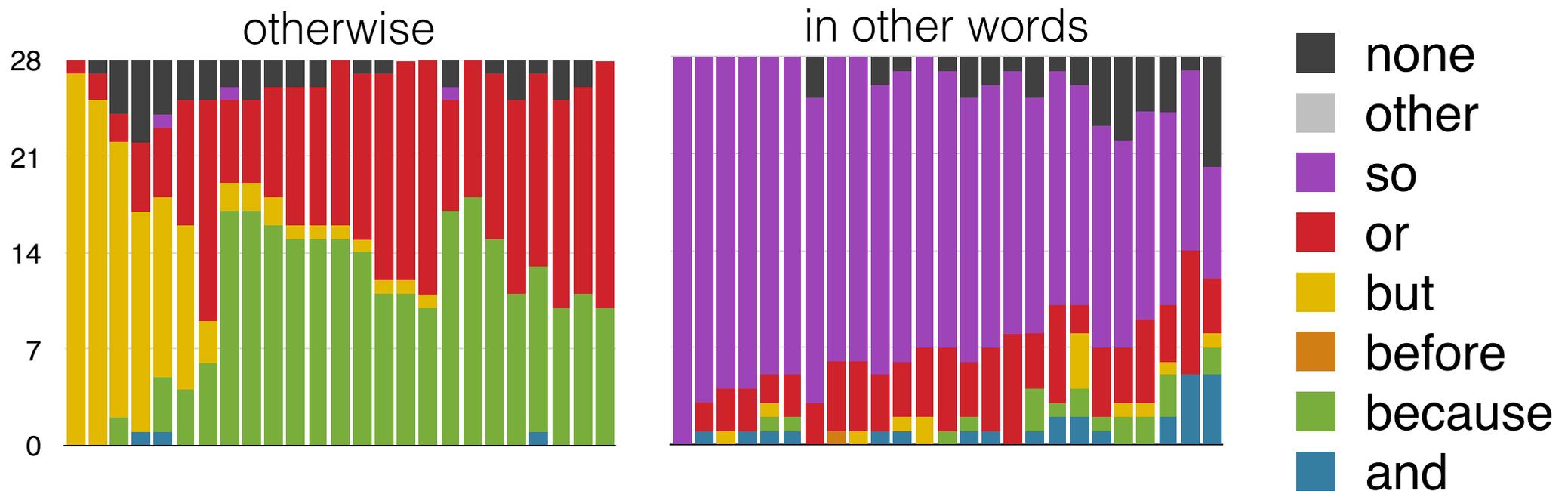→ passage requires causal reasoning (BECAUSE)

Unfortunately, nearly 75,000 acres of tropical forest are converted or deforested every day _____ in other words an area the size of Central Park disappears every 16 minutes.

→ *in other words* encodes 'otherness' (OR)
→ reformulation conveys consequence (SO)

e.g., Adverbials that encode 'alternative'

▶ Adverbial meaning: *otherwise* and *in other words* license OR

▶ Additional pragmatic inference: Passage content licenses BECAUSE in some cases, SO in others

▶ What determines coherence relation with *otherwise*?

Different underlying pragmatic logic

argumentation

Proper placement of the testing device is an important issue _____ otherwise the test results will be inaccurate.

"a reason to place the test properly is to avoid inaccuracy"

enumeration

A baked potato, plonked on a side plate with sour cream flecked with chives, is the perfect accompaniment _____ otherwise you could serve a green salad and some good country bread.

"there's more than one option for a side: potato or salad"

exception

Mr. Lurie and Mr. Jarmusch actually catch a shark, a thrashing 10-footer _____ otherwise the action is light.

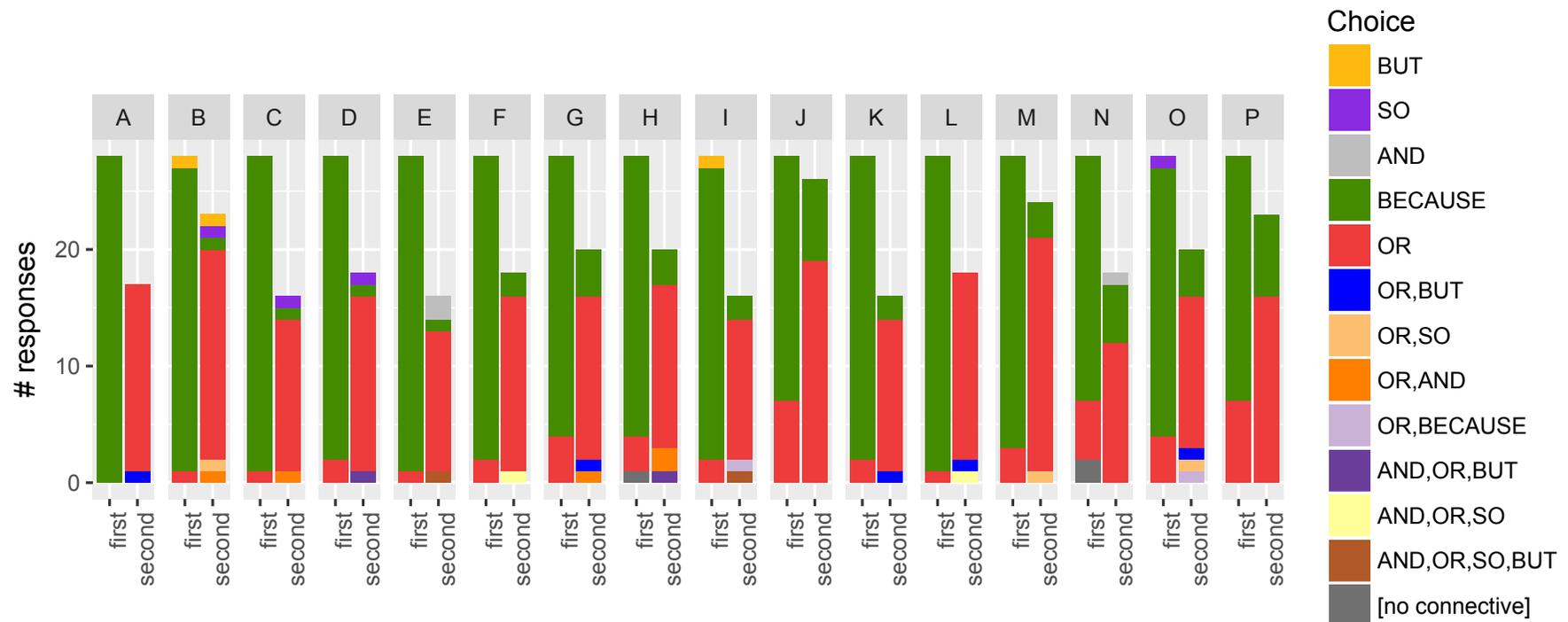"shark catching is a special case; generally action is light"

New study: Insert conjunction(s)

▸ **Materials:** 48 passages with *otherwise* (16 argumentation, 16 exception, 16 enumeration)

▸ **Participants:** 28 participants

▸ **Task 1:** Find best conjunction(s) for meaning of connection

▸ **Task 2:** Find paraphrase of that meaning

Results

argumentation **Proper placement of the testing device is an important issue _____ otherwise the test results will be inaccurate.**
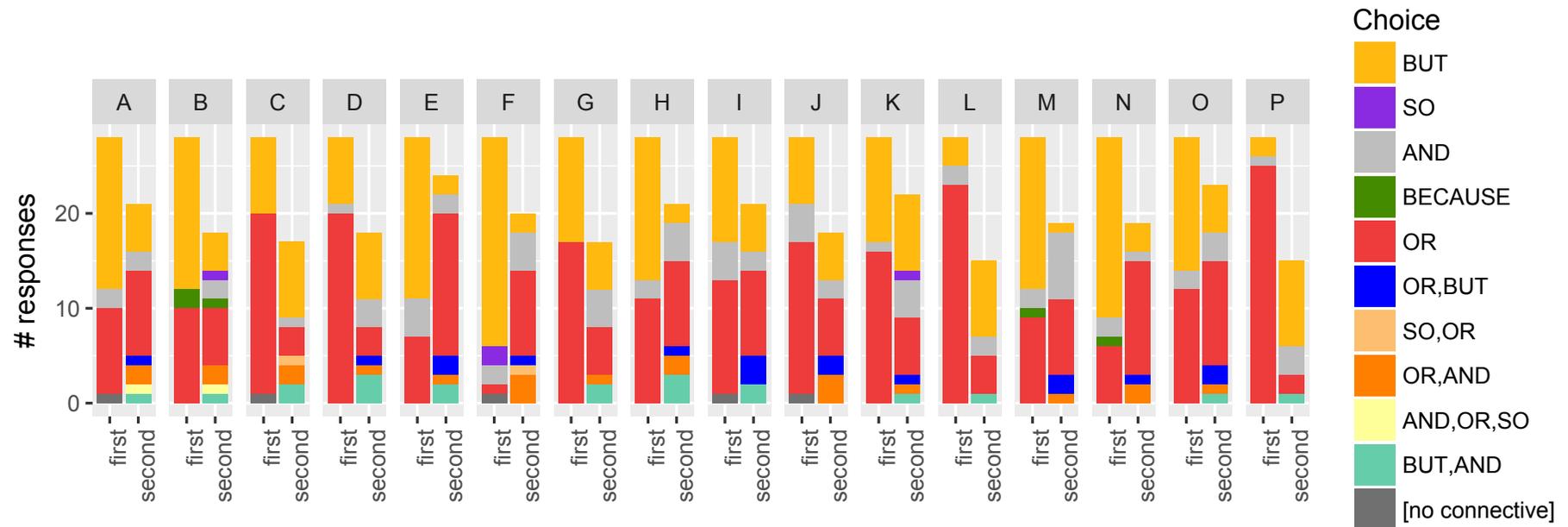
→ Prediction confirmed: OR/BECAUSE [99%]

Results

enumeration

A baked potato, plonked on a side plate with sour cream flecked with chives, is the perfect accompaniment ____ otherwise you could serve a green salad and some good country bread.
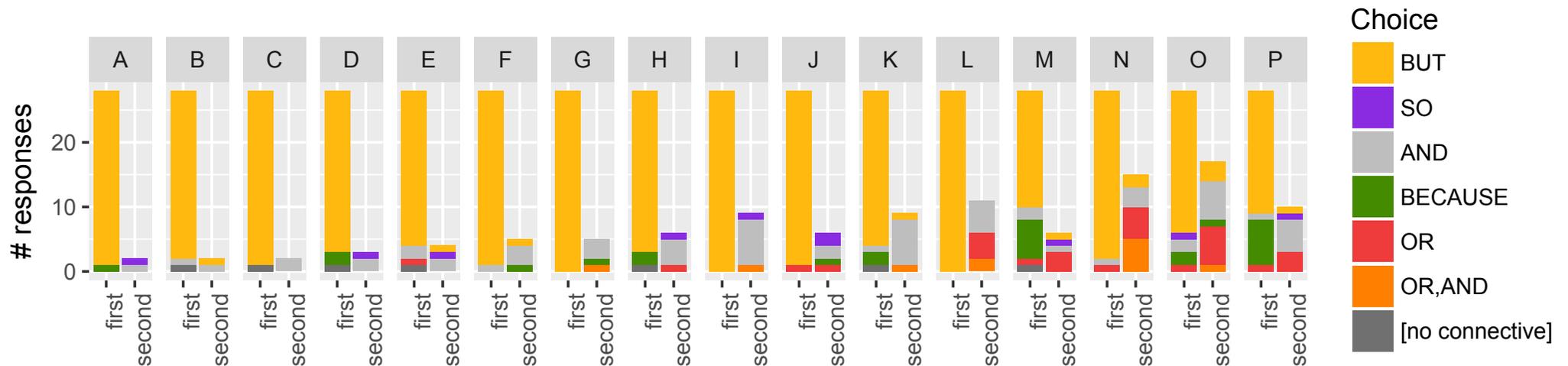
→ Prediction confirmed: OR/BUT [98%]

Results

exception

Mr. Lurie and Mr. Jarmusch actually catch a shark, a thrashing 10-footer _____ otherwise the action is light.

→ Prediction confirmed: BUT [92%]

Summary: Choosing among alternatives

▶ Multiple co-occurring relations

> **It's too far to walk. Let's take the bus.**

▶ Inference even with explicit cues

> **It's too far to walk.** [result] **Instead let's take the bus.**

> **Better to take the bus** [reason] **or otherwise you'll have to walk.**

▶ Informative disagreement

Thanks!

Paraphrase results

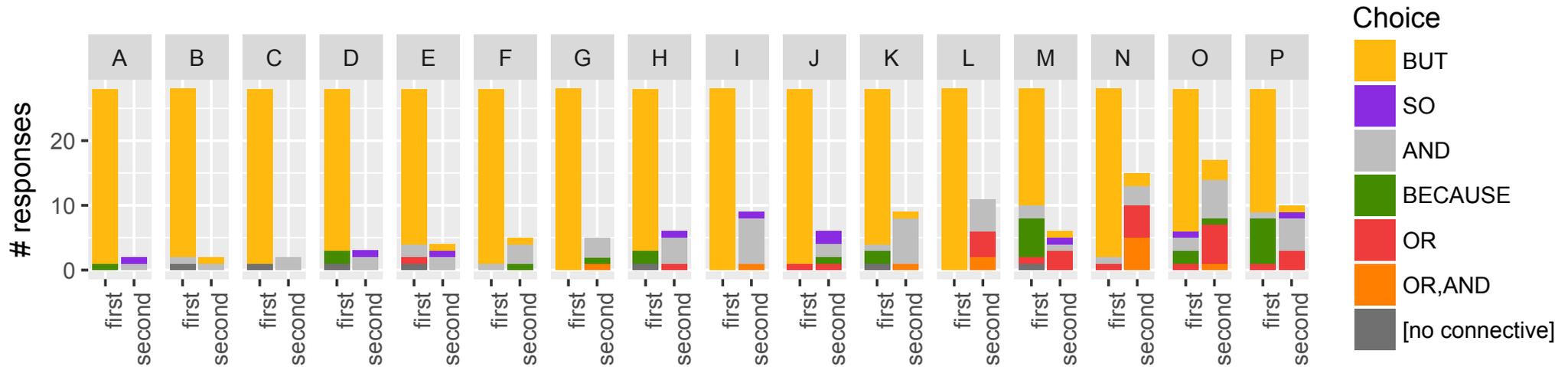| Researcher	Participant			
	Argumentation	ENUMERATION	Exception	Multiple
ARGUMENTATION	401 (91.5%)	4	25	18
ENUMERATION	23	364 (81.4%)	46	14
EXCEPTION	21	29	393 (87.7%)	5

Table 1: Researcher labels assigned to *otherwise* passages vs. labels implied by participant paraphrases

Proper placement of the testing device is an important issue _____ otherwise the test results will be inaccurate.

"a reason to place the test properly is to avoid inaccuracy"

"inaccurate test results are a special case; generally proper placement is important"

Results

M

  Democrats insist that the poor should be the priority, and that tax relief should be directed at them _____ otherwise they lack a cogent vision of the needs of a new economy.

P

  He said that the proposed bill would give states more flexibility in deciding whether they wanted to use the Federal money for outright grants to municipalities or to set up loan programs _____ otherwise it left last fall's Congressional legislation unchanged.

Results: "In other words"

Unfortunately, nearly 75,000 acres of tropical forest are converted or deforested every day _____ in other words an area the size of Central Park disappears every 16 minutes.

→ *in other words* encodes 'otherness' (OR)
→ reformulation conveys consequence (SO)

Unfortunately, nearly 75,000 acres of tropical forest are converted or deforested every day. *I don't know where I heard that* _____ in other words an area the size of Central Park disappears every 16 minutes.

→ intervening material blocks adjacency preferred for OR/SO, allowing more uses of BUT to emerge