

Why are you telling me this?

Relevance & informativity in language processing

Hannah Rohde
University of Edinburgh

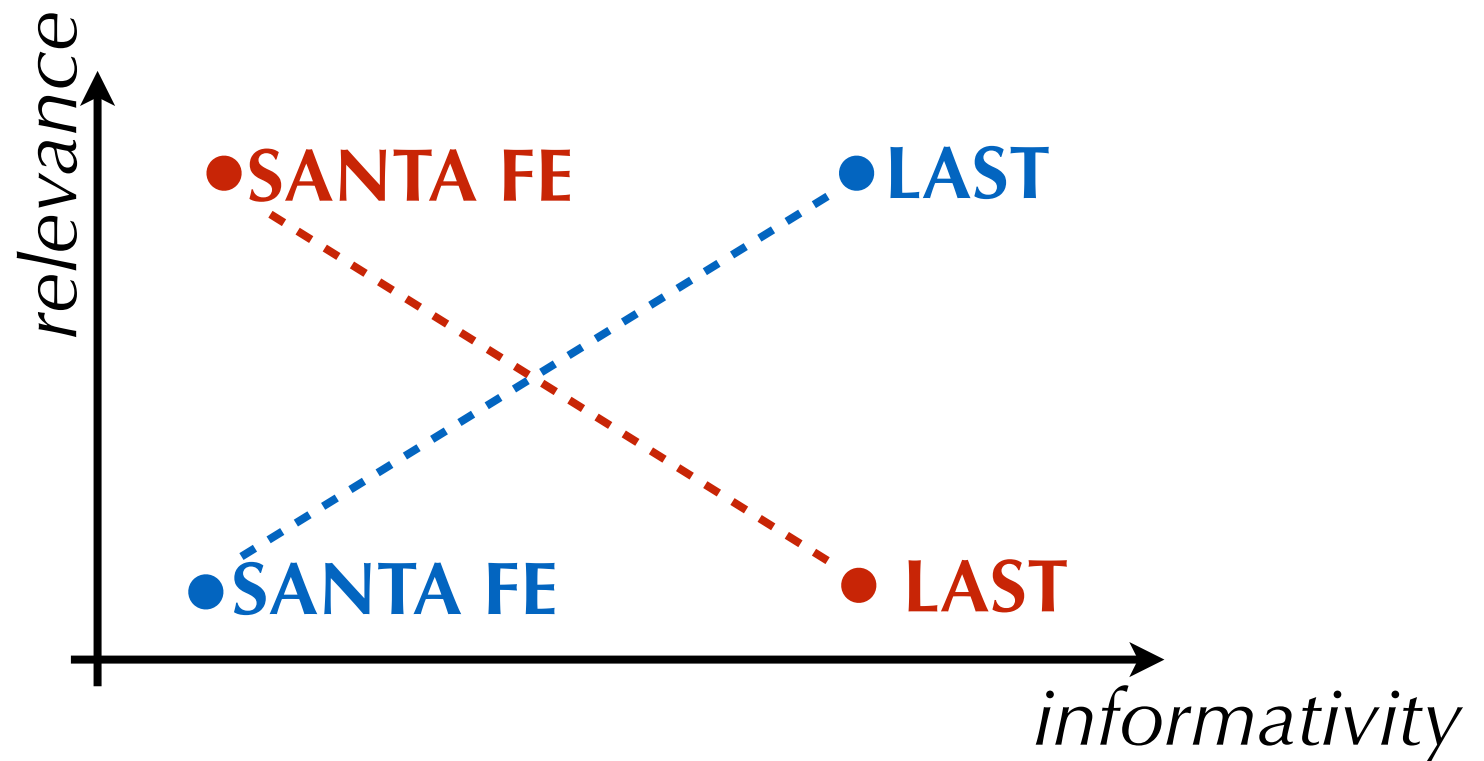
(1) The Santa Fe Trail goes to Santa Fe.

(2) Today is the last day of your life.

**(2) is more informative more than (1)
But which one is more relevant?**

Informativity & relevance

- Where does the Santa Fe Trail go?
- How long will I live?



→ An utterance's status depends on its ability to reduce uncertainty and address a question under discussion (QUD)

Orthogonal?

- ▶ If informativity is about addressing an open QUD, then **informativity = relevance**.
- ▶ But if informative utterances yield belief updates, then **informativity \neq relevance**

- ▶ Informative utterances update prior probabilities

Today is the last day of your life.

For breakfast, I ate twenty bananas.

For breakfast, I ate one banana.

For breakfast, I ate one yellow banana.

The Santa Fe Trail goes to Santa Fe.

- ▶ Relevant utterances address probable QUDs

This talk: How do we infer relevance and how do we process (un)informative information?

▶ Relevance relations

- ▶ Where to look? [multiple concurrent relations]
- ▶ What to listen for? [focus intonation]
- ▶ What cues? [adverbials, verbs, segment properties]
- ▶ Where else to look? [relative clauses]

→ **Don't miss available relations or (machine-identifiable) cues**

▶ Informativity

- ▶ Redundant facts ["dozen cookies...12"]
- ▶ Redundant visual cues [REG]

→ **Not all redundant information is irrelevant**

Discourse coherence relations

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.

(because)

^ Otherwise

you'll be left with soggy cupcakes.



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

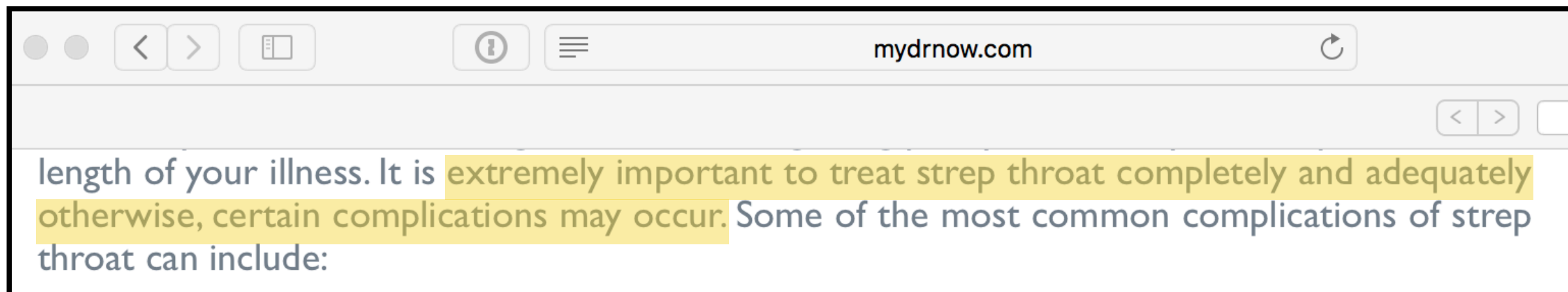
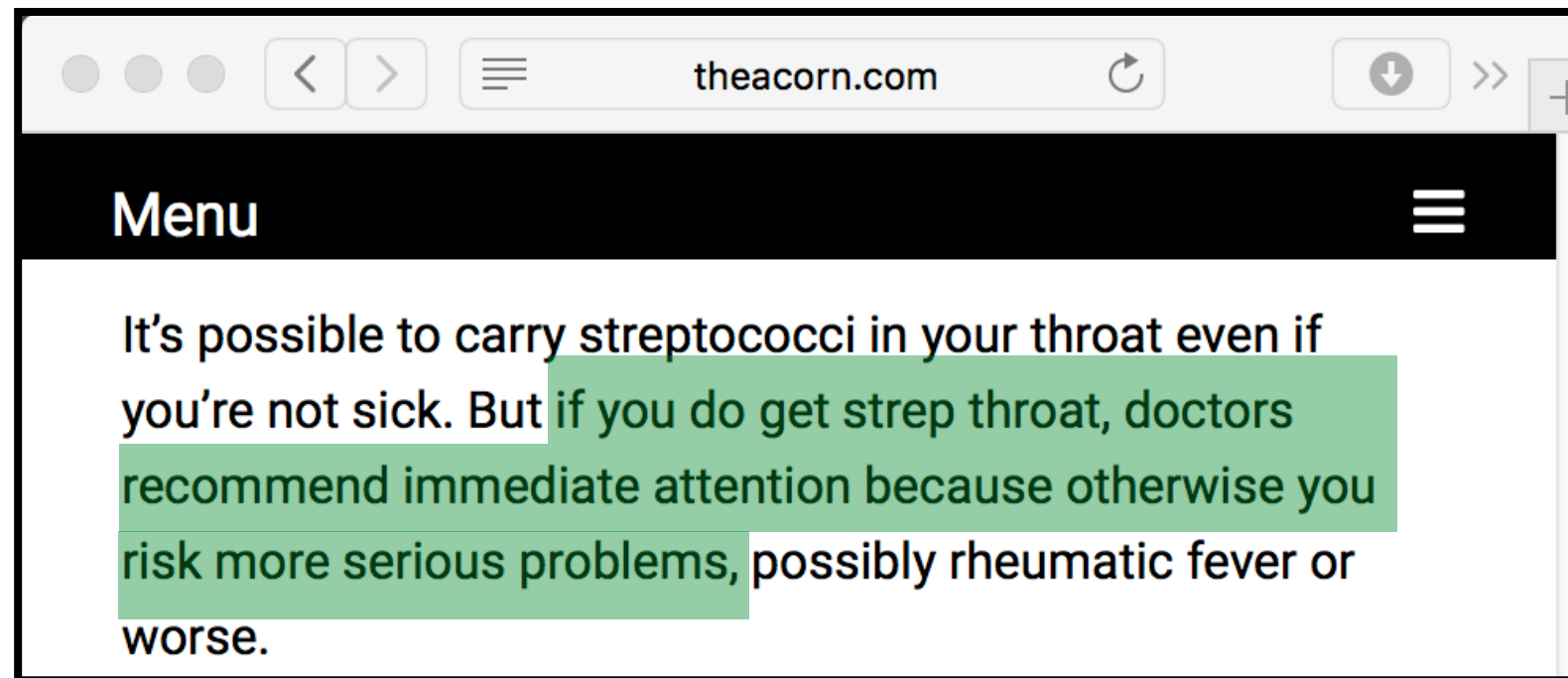
→ Inference occurs alongside overt connectives.



Coherence relations in NLP

► Question-answering

Query: “why treat strep throat?”

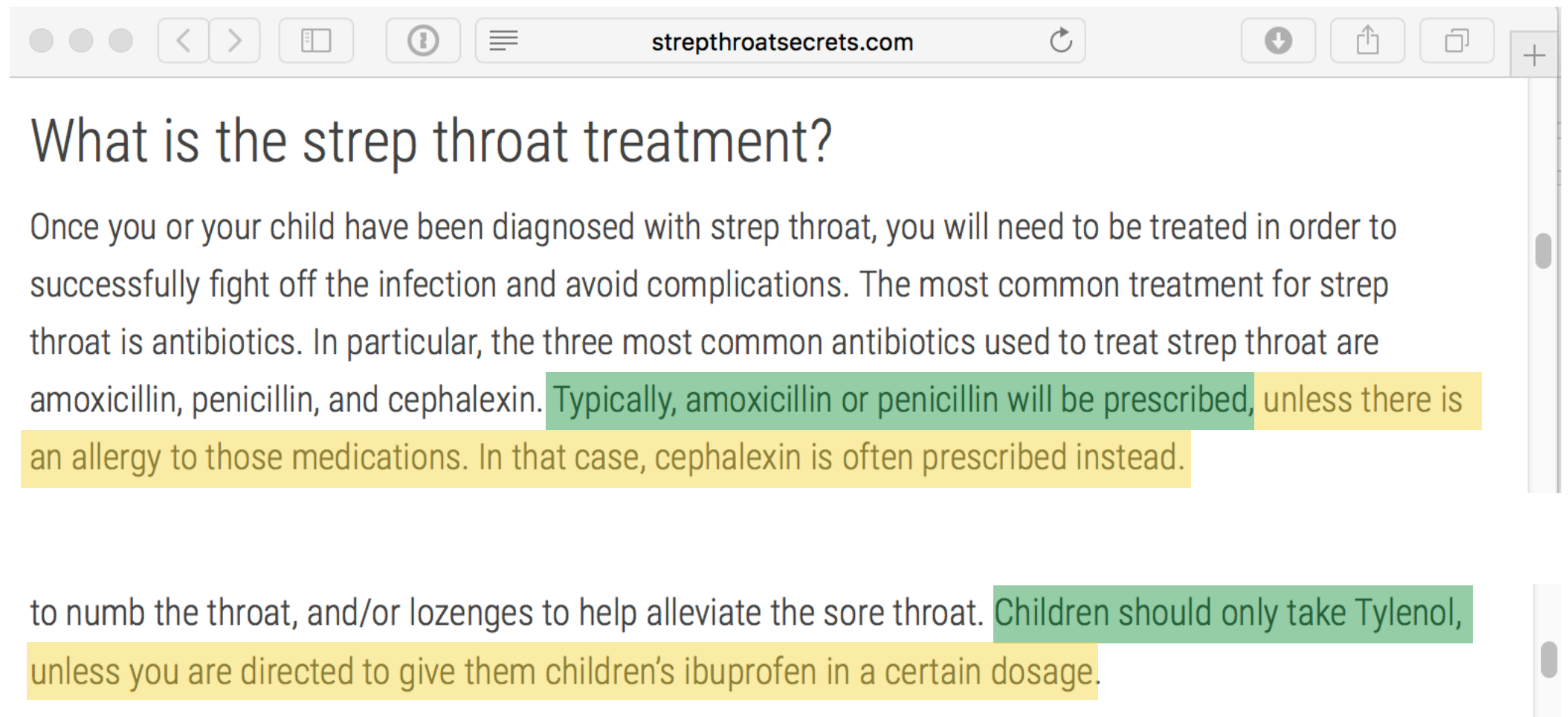


Coherence relations in NLP

► Question-answering

Query: “why treat strep throat?”

Query: “how to treat strep throat?”



The screenshot shows a web browser window with the address bar displaying 'strepthroatsecrets.com'. The main content area has the heading 'What is the strep throat treatment?'. Below the heading is a paragraph of text. The text is highlighted in yellow, and the phrase 'Typically, amoxicillin or penicillin will be prescribed, unless there is an allergy to those medications. In that case, cephalexin is often prescribed instead.' is highlighted in green. Below this paragraph is another sentence: '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.' This sentence is also highlighted in yellow.

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 NLP

▶ 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

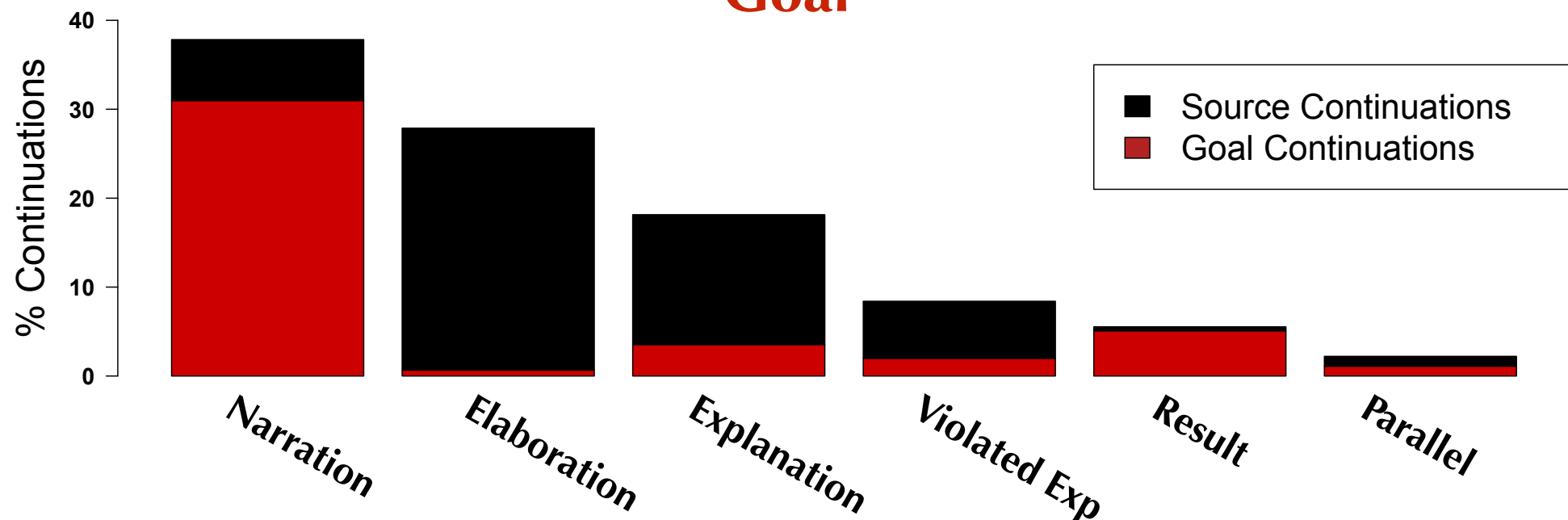
Coherence relations in NLP

- ▶ Question-answering
- ▶ Text generation, automatic summarisation:
What to make explicit to sound natural?
- ▶ Coreference resolution
Best antecedent may vary across coherence relations.

John handed a book to Bob. He then thanked John for the book.

(Kehler & Rohde, 2013)

Goal



Coherence relations in NLP

- ▶ Question-answering
- ▶ Text generation, automatic summarisation:
What to make explicit to sound natural?
- ▶ Coreference resolution
Best antecedent may vary across coherence relations.
- ▶ Given this utility,
 - ▶ large-scale annotated resources
 - ▶ discourse parsing tasks

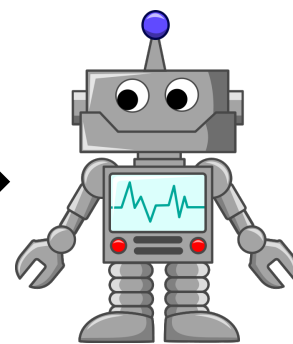
Assumption: implicit *or* explicit relations

 **Cover cakes with this frosting**
Otherwise you'll be left with soggy cupcakes.



- ▶ How widespread is inference alongside explicit connectives?
- ▶ How much variation across adverbials/passages?

- ▶ If deterministic →



- ▶ If not →



Expt1: Conjunction insertion task

ConnText University of Edinburgh

Trial Show Instructions

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

Conjunction:

☐ because

☐ or

☐ but

☐ so

☐ and

☐ none at all

☐ 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

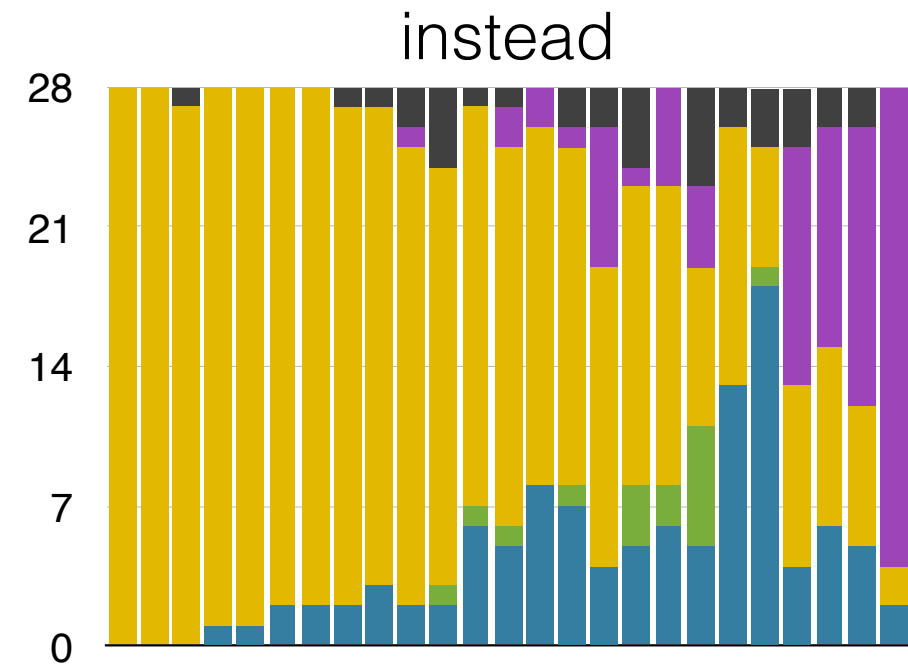
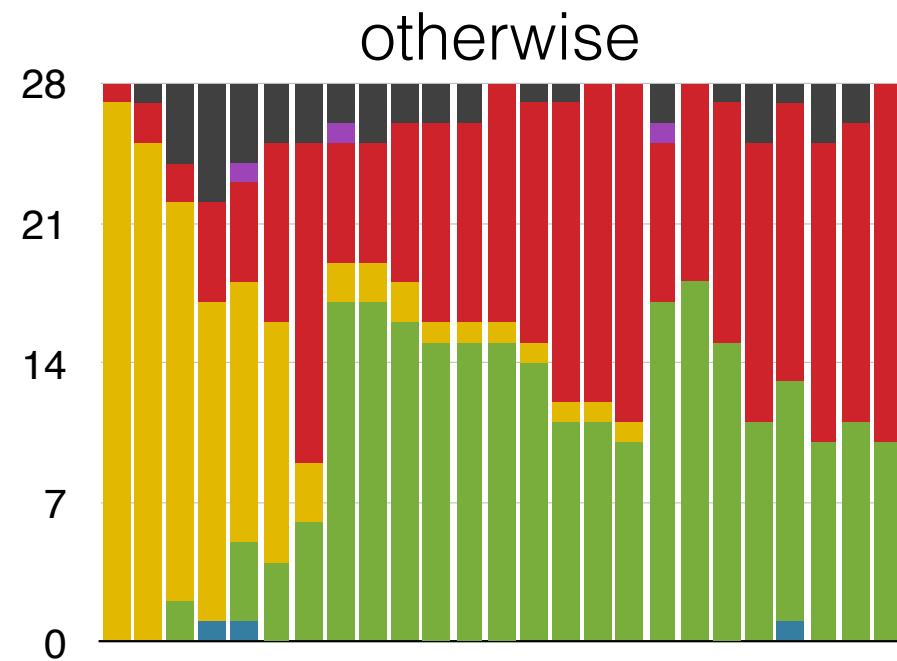
Materials: Passages from NYTimes, half implicit, half explicit

(Rohde, Dickinson, Schneider, Louis, & Webber, 2017, IWCS; see also Scholman et al. 2016)

accordingly	for instance	in turn	overall
actually	for one thing	indeed	previously
additionally	furthermore	instead	really
after all	hence	later	similarly
afterwards	however	likewise	specifically
alternatively	in addition	meanwhile	still
as a result	in contrast	moreover	then
consequently	in fact	nevertheless	thereafter
earlier	in general	nonetheless	therefore
ever since	in other words	of course	thus
finally	in particular	on the one hand	ultimately
first	in that case	on the other hand	what's more
first of all	in the end	otherwise	yet
for example			

- 50+ adverbials, each in 50+ passages, 28 people/passage
- 70,000+ judgments

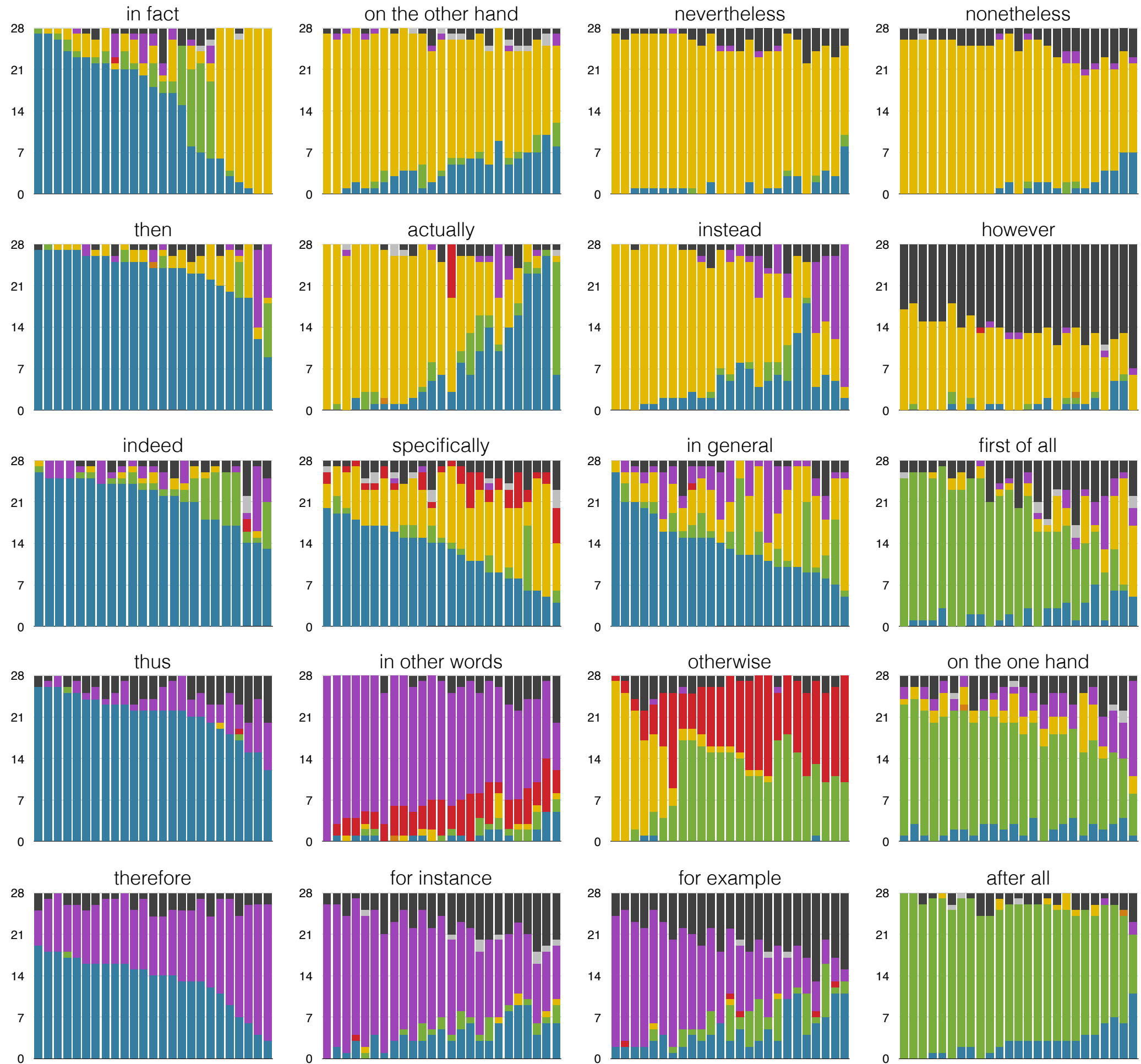
■ and
 ■ because
 ■ before
 ■ but
 ■ or
 ■ so
 ■ other
 ■ none



→ Not uniform across adverbials

→ Not uniform within adverbials

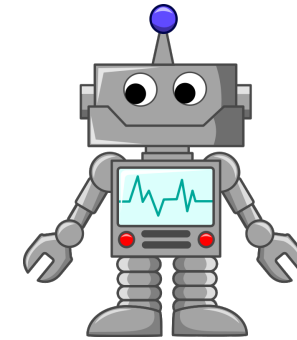
and because before but or so other none



Multiple concurrent relations

- ▶ Inference is widespread alongside explicit connectives and varies across adverbials and passages

- ▶ Pockets of systematicity →



- ▶ Not deterministic →

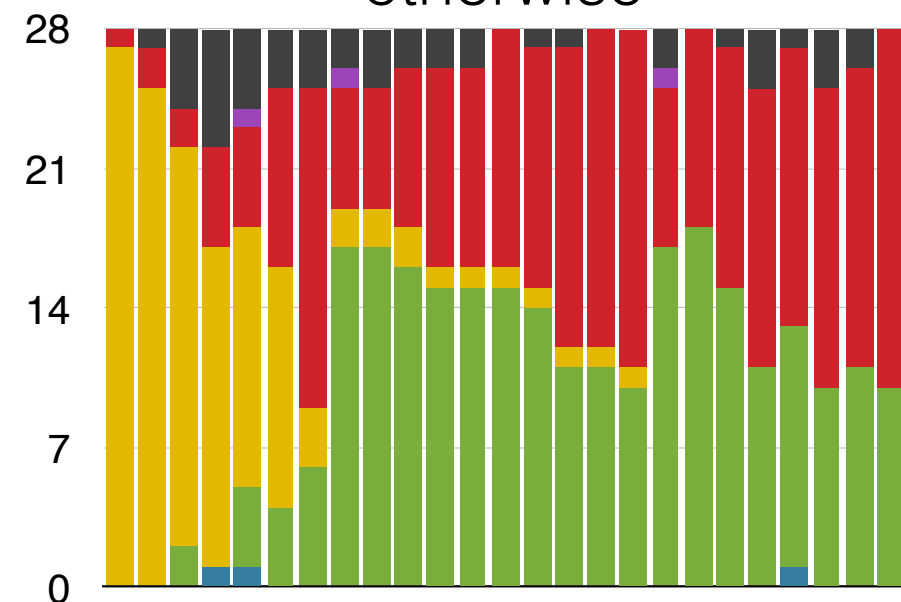


- ▶ What if humans disagree?

Inter-annotator disagreement

but or because

otherwise



You got to be nice to them // _____ otherwise they're not going to be nice to you.

Author=OR

14 Participants=OR

13 Participants=BECAUSE

1 Participant=NONE

- ▶ Not evidence of mistakes or ambiguity
- ▶ Improbable combinations, but both valid
- ▶ Multiple concurrent discourse relations

Expt2: Which conjunction(s)?

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 the side with sour cream and chives, is the perfect accompaniment _____ otherwise you could serve a green salad and some good country bread.

"there are two choices for a side: potato or salad"

"a reason to have a potato is to avoid a salad"

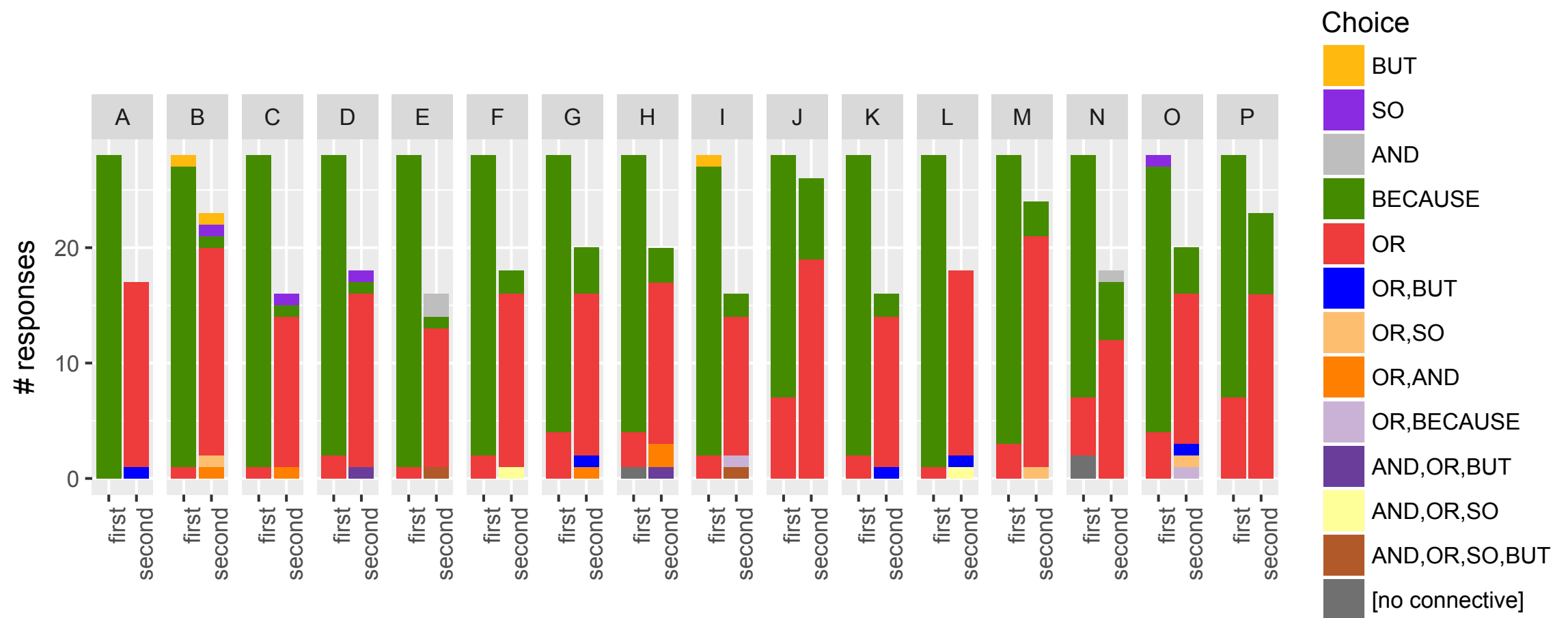
exception Mr. Lurie and Mr. Jarmusch actually catch a shark, a thrashing 10-footer _____ otherwise the action is light.

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

"there are two choices for a film: sharks or light action"

Results: argumentation passages

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



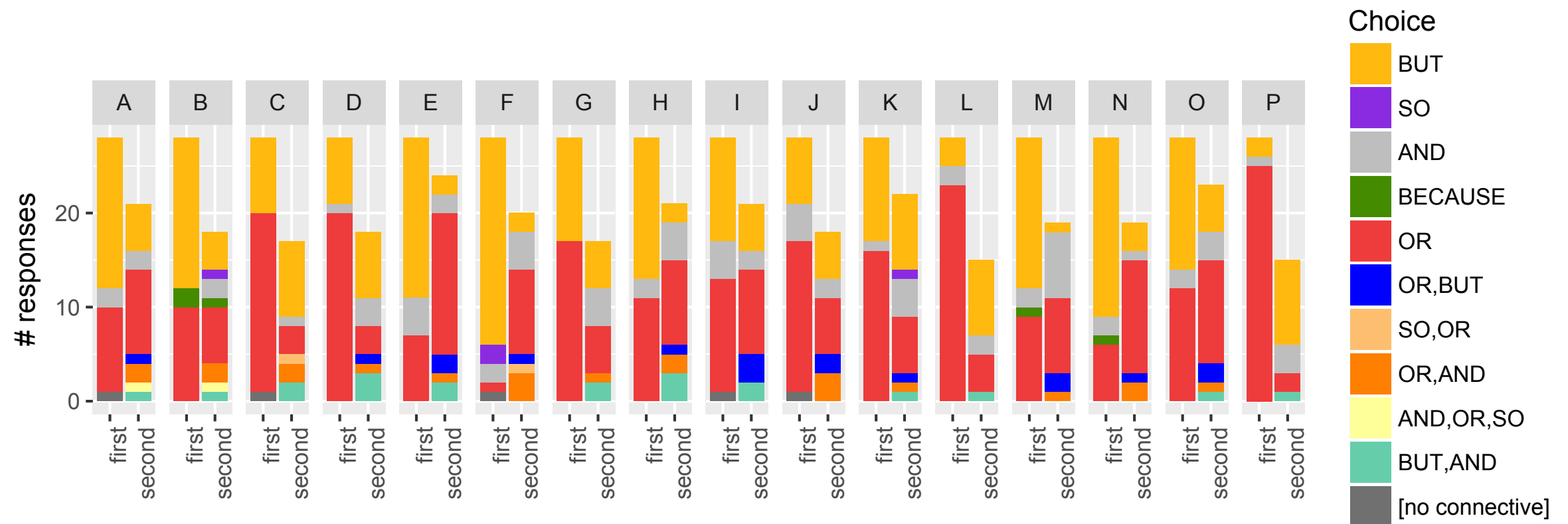
→ Confirmed: BECAUSE & OR

→ Cue: Segment 2 contains undesirable outcome

Results: enumeration passages

enumeration

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



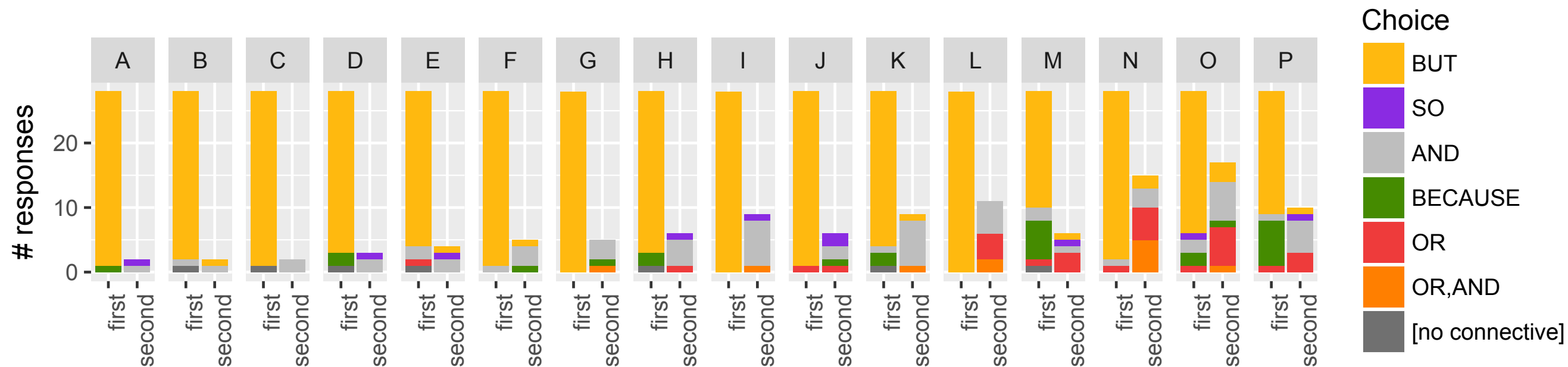
→ Confirmed: BUT & OR

→ Cue: Segments 1 & 2 list equal alternatives

Results: exception passages

exception

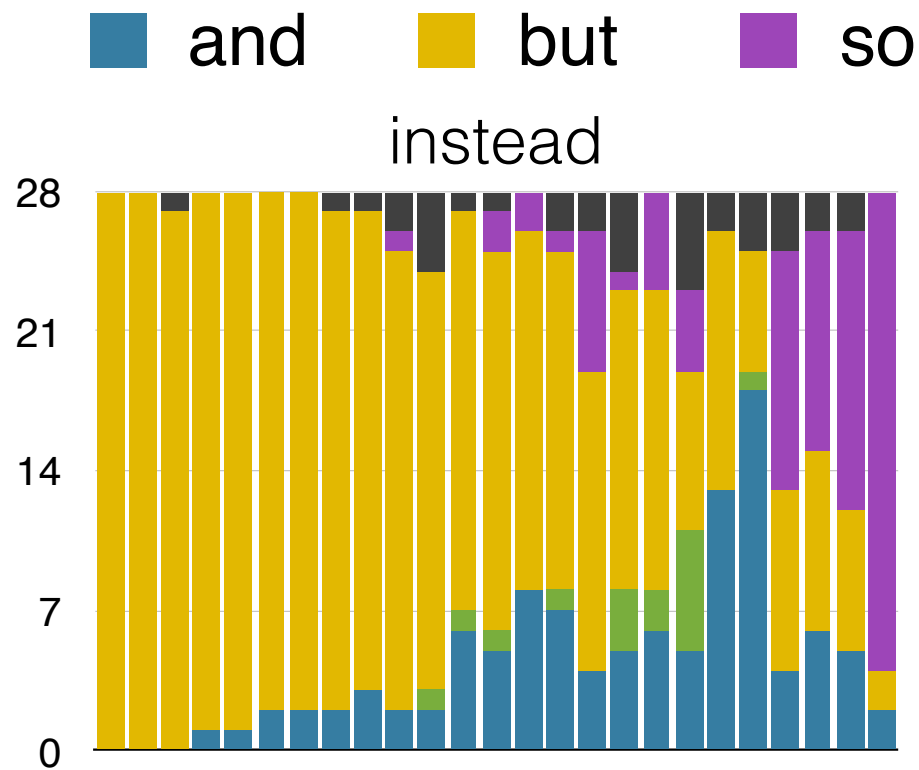
Mr. Lurie and Mr. Jarmusch actually catch a shark, a thrashing 10-footer _____ otherwise the action is light.



→ Confirmed: BUT only

→ Cue: Segment2 describes a generalisation

Expt3: Different passage logic (*instead*)



Contrast

I was really hoping for a promotion
// _____ instead I got fired.

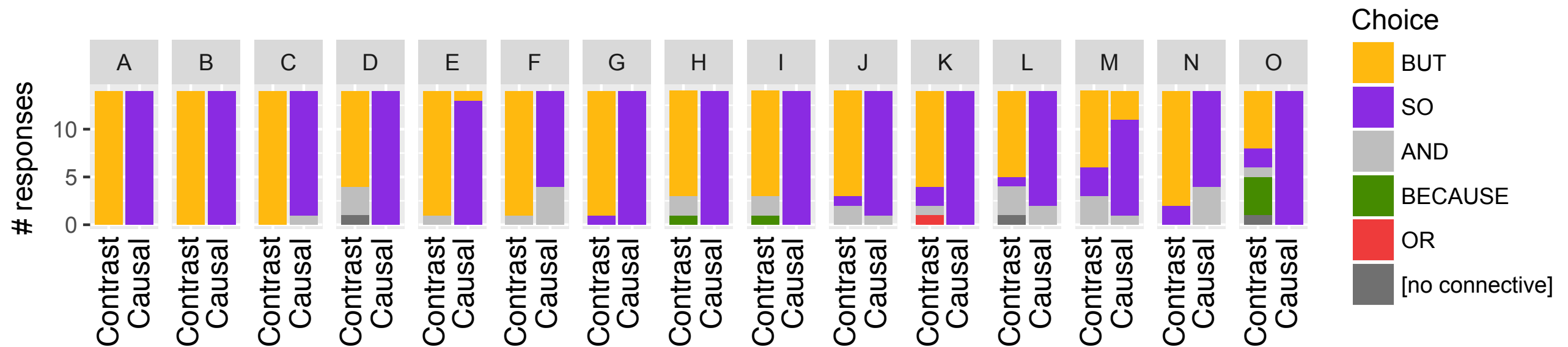
Verb cue: *hoping/wanting/planning* → BUT

Causal

I was too pushy when I applied for a promotion
// _____ instead I got fired.

Downward entailing operator: *too+[adj]* → SO

Results: contrast vs causal passages



- Confirmed: segment properties predict BUT vs SO
- Use of human studies to identify available interpretations and relevant cues (see use of corpus annotations of relation signals: Taboada & Das 2013; Zeldes 2018)

This talk: How do we infer relevance and how do we process (un)informative information?

- ▶ Relevance relations
 - ▶ Where to look? [multiple concurrent relations]
 - ▶ What to listen for? [focus intonation]
 - ▶ What cues? [adverbials, verbs, segment properties]
 - ▶ Where else to look? [relative clauses]
- ▶ Informativity
 - ▶ Redundant facts ["dozen cookies...12"]
 - ▶ Redundant visual cues [REG]

Recovering QUDs

- ▶ Intonation can signal the question under discussion
(Büring, 2004; Most & Saltz, 1979; Roberts, 1996)

THE PITCHER threw the ball.

→ “Who threw the ball?”

The pitcher threw **THE BALL**.

→ “What did the pitcher throw?”

- ▶ Coherence relations can be understood as QUDs

Mary congratulated Sue. She won the race.

Why

- ▶ But is there a “Why” intonation?

Expt4: What to listen for?

Charles congratulated Simon. He ...

- ▶ Implicit causality verbs: *congratulate/scold/admire/...*
 - ▶ Create expectation for answer to “Why?”
 - ▶ For IC2 verb, causally implicated referent = 2nd NP
 - ▶ Different QUDs → different interpretations of pronoun
 - ▶ What did Charles do and why? [*because Simon...*]
 - ▶ What all did Charles do? [*and also Charles...*]
- Causal relation favors causally implicated Simon
- Focus intonation may signal a parallel relation, reducing bias to Simon

Expt4: What to listen for?

N=75, vary intonation, only NP2-biased verbs

IC Intro Charles congratulated Simon.

- focus He had criticized Stephanie.

+focus He had **CRITICIZED STEPHANIE.**

Task: Who criticized Stephanie? _____

- ▶ Replicate known implicit causality bias: 65% NP2
- ▶ Reduce that bias with +focus: 59% NP2
- ▶ Intonation guides relation, relation guides coreference

Other cues to upcoming relations

- ▶ Adverbials can establish long-distance dependencies
(Scholman, Rohde, & Demberg, 2017)

[On the one hand ...] On the other hand ...

- ▶ Verb class guides expectations in story continuations
(Kehler, Kertz, Rohde & Elman 2008; Rohde & Horton 2014)

[implicit causality] Mary congratulated Sue. ... → explanation

[transfer] Mary handed a book to Sue. ... → narration

- ▶ Event structure constrains upcoming relations
(Kehler, Kertz, Rohde & Elman 2008; Rohde & Horton 2014)

[imperfective] Mary was handing a book to Sue. ... → elaboration/
violated expectation

This talk: How do we infer relevance and how do we process (un)informative information?

- ▶ Relevance relations
 - ▶ Where to look? [multiple concurrent relations]
 - ▶ What to listen for? [focus intonation]
 - ▶ What cues? [adverbials, verbs, segment properties]
 - ▶ Where else to look? [relative clauses]
- ▶ Informativity
 - ▶ Redundant facts ["dozen cookies...12"]
 - ▶ Redundant visual cues [REG]

What can a relative clause do?

- ▶ CLAIM: Only some relative clauses serve as discourse segments (Mann & Thompson 1988; Reese et al. 2007; Sanders & van Wijk 1996; Verhagen 2001)

- ▶ ✗ Restrictive RCs — only aid in establishing reference

Mary congratulated the guy who lives next door.

- ▶ ✓ Non-restrictive RCs — can enter into relevance relations

Mary congratulated Bob, who won the lottery. [reason]

- ▶ What about restrictive RCs with simultaneous relevance?

➡ Mary congratulated the guy who won the lottery. [reason]

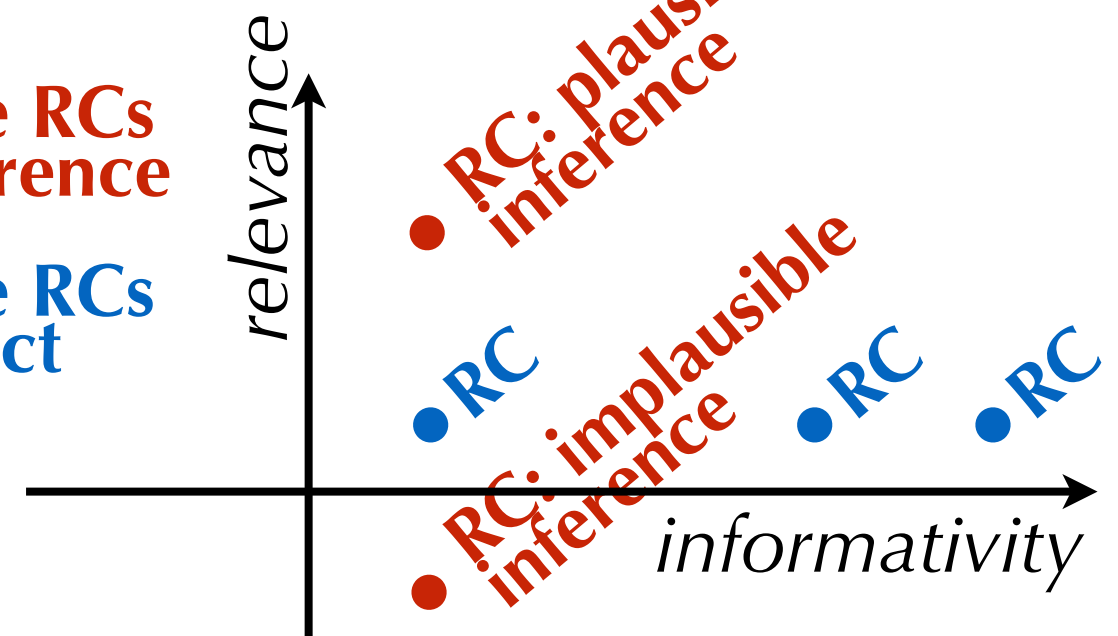
Expt5: Where else to look?

Self-paced reading, N=52, vary matrix verb

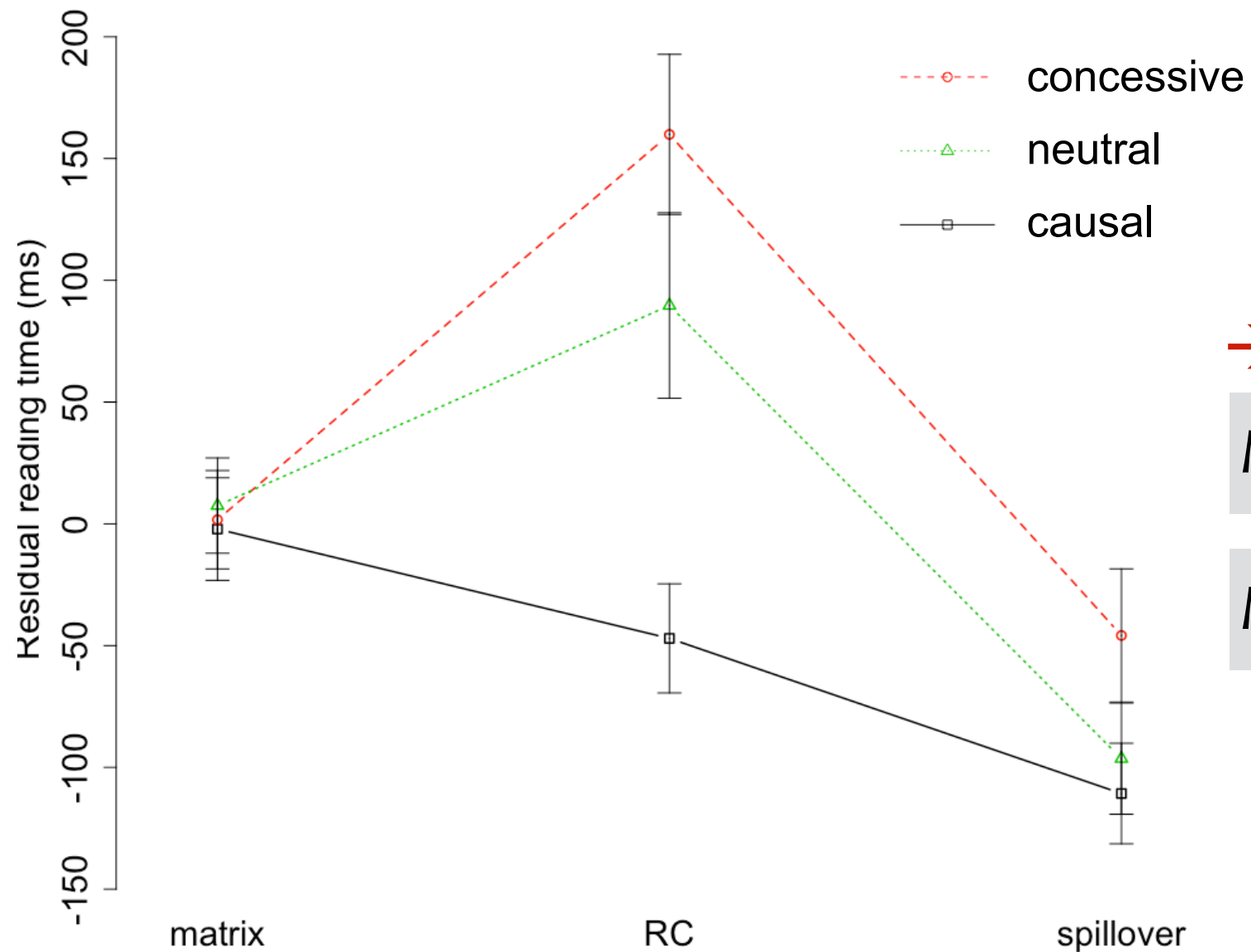
Intro	Jenny walked through the hallway to check on the daily goings-on around the office.
causal RC	She <u>congratulated</u> the guy who made lots of money for the company.
neutral RC	She <u>joked</u> with the guy who made lots of money for the company.
concessive RC	She <u>fired</u> the guy who made lots of money for the company.
Wrap-up	She arrived at the conference room in time for her next meeting.

● Restrictive RCs allow inference

● Restrictive RCs only restrict reference



Expt5: Where to look?



→ Where else?

Mary scolded the lazy student.

Mary congratulated the winner.

→ Infer relevance of RC to matrix clause during real-time processing

This talk: How do we infer relevance and how do we process (un)informative information?

▶ Relevance relations

- ▶ Where to look? [multiple concurrent relations]
- ▶ What to listen for? [focus intonation]
- ▶ What cues? [adverbials, verbs, segment properties]
- ▶ Where else to look? [relative clauses]

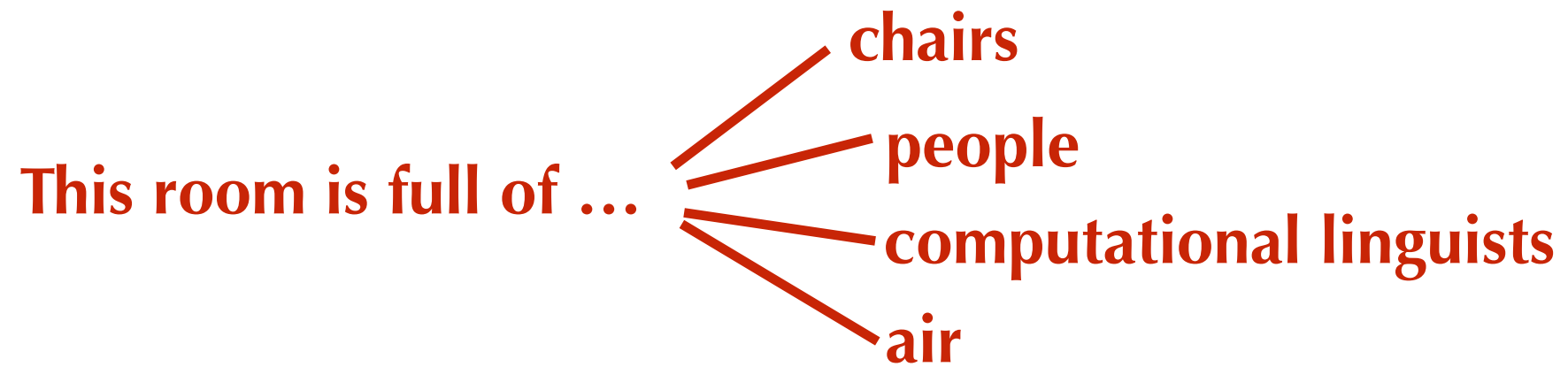
→ **Don't miss available relations or (machine-identifiable) cues**

▶ Informativity

- ▶ Redundant facts ["dozen cookies...12"]
- ▶ Redundant visual cues [REG]

bothering

Why are you telling me this?



- ▶ “air” is very likely to be true, but it’s uninformative
- ▶ nonetheless, probable stuff often appears easy to process

Predictability in psycholinguistic studies



The Dutch trains are ...

- yellow
- white
- sour

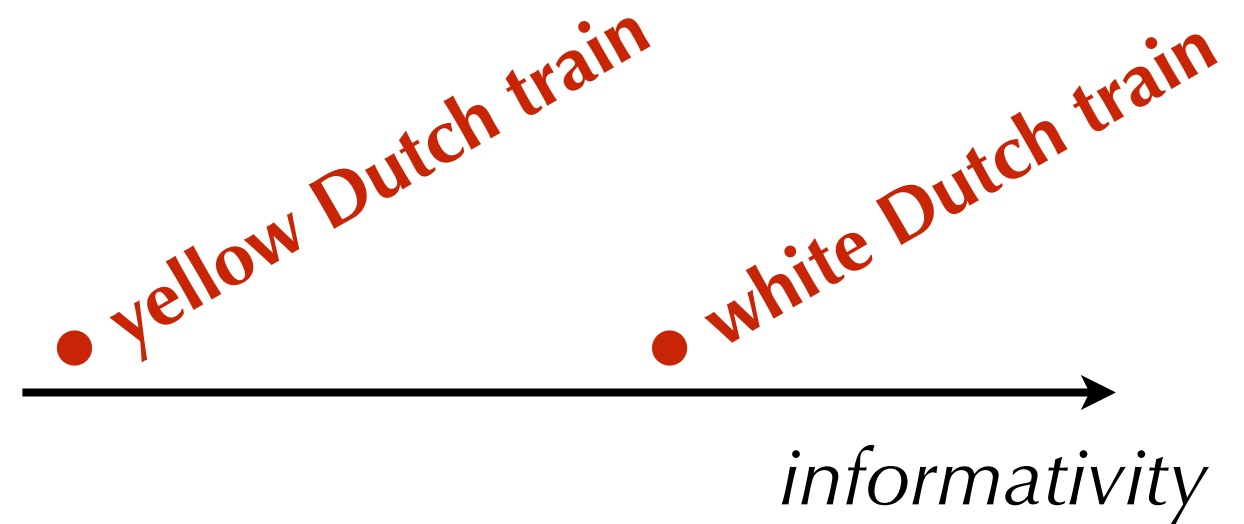
- ▶ Improbable words yield more surprisal than probable ones (Hagoort et al. 2004)
sour > white > yellow
- ▶ Uninformative material is fine in the lab (cf. Kravtchenko & Demberg 2015)

Uninformativity outside the lab?



This is not how we use language!
(Stalnaker 1973; Grice 1975)

Expt6: Redundancy in reading



Self-paced reading of text messages,
N=214, IbexFarm, vary informativity

[informative] promised a dozen cookies
... baked 5

[duh] promised a dozen cookies
... baked 12

“yellow Dutch train”

p(situation) is high
p(utterance | situation) is low

“white Dutch train”

unlikely, but interesting!

Expt6: Redundancy in reading

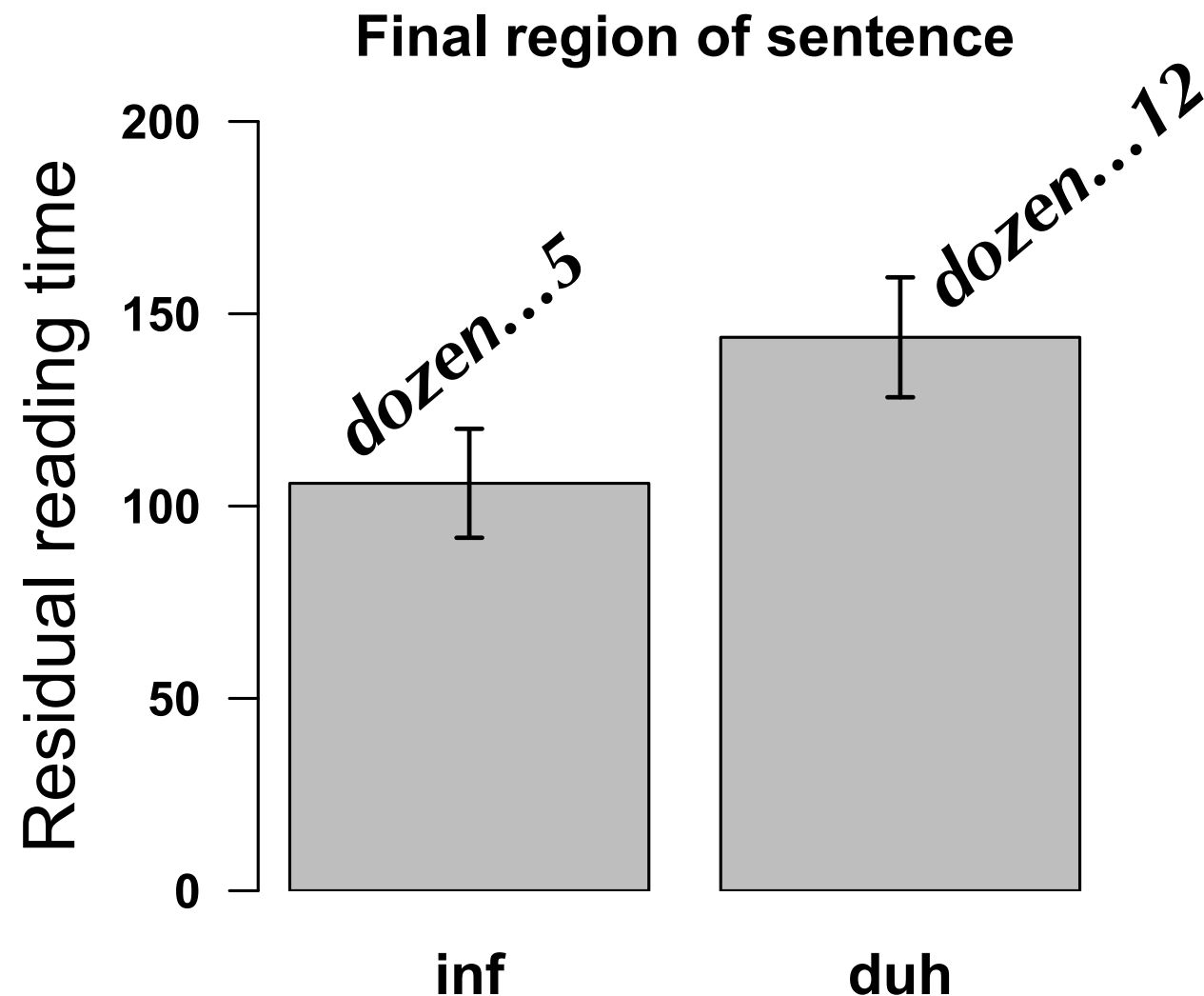


Expt6: Redundancy in reading

_____ year

scenario	inf	duh
bake a dozen cookies	5	12
invite 5 people	12	5
toddler's age	25	5
25-year-old's age	5	25
cost of one sock	\$150	\$2
cost of a leather jacket	\$2	\$150
cost of a headband	\$200	\$10
cost of a Versace scarf	\$10	\$200

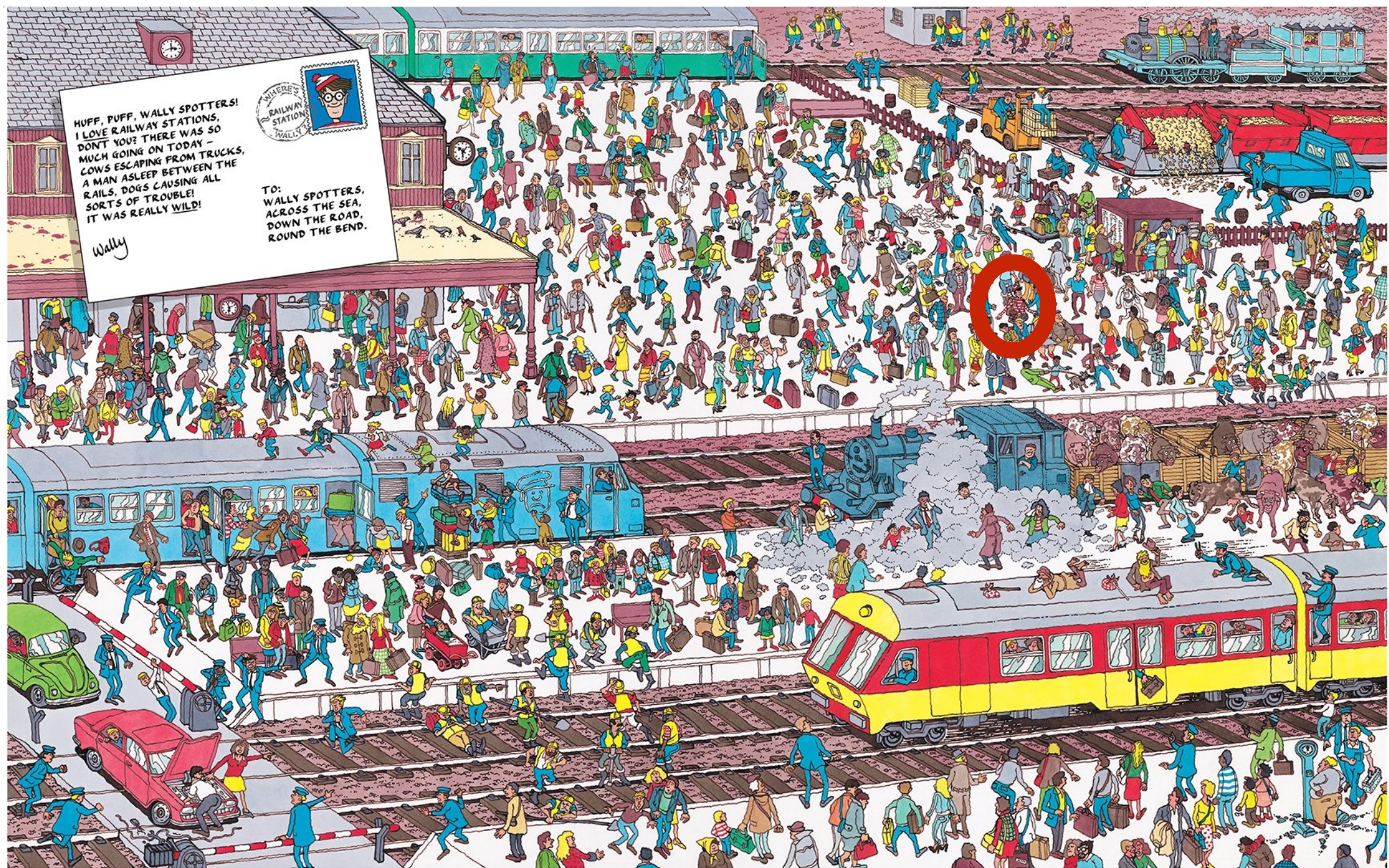
Expt6: Redundancy in reading



- People expected to be surprised, unlike in prior lab studies
- Uninformative messages are hard, even if content is predictable

This talk: How do we infer relevance and how do we process (un)informative information?

- ▶ Relevance relations
 - ▶ Where to look? [multiple concurrent relations]
 - ▶ What to listen for? [focus intonation]
 - ▶ What cues? [adverbials, segment properties, verbs]
 - ▶ Where else to look? [relative clauses]
- ▶ Informativity
 - ▶ Redundant facts ["dozen cookies...12"]
 - ▶ Redundant visual cues [REG]



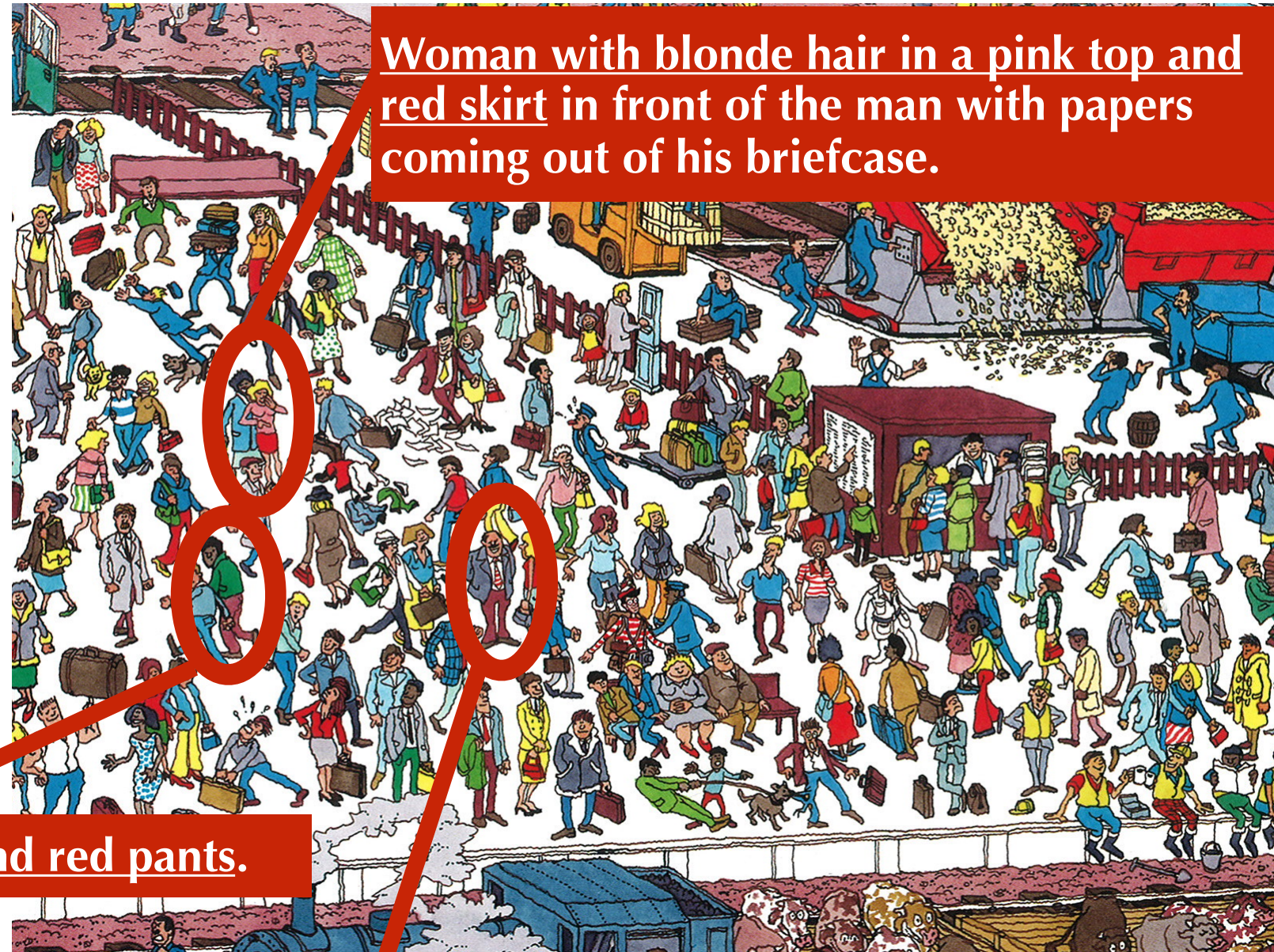
Find the guy in glasses wearing the red and white striped hat and shirt
who is behind the bench
and above the cows
that are on the train.

redundant

Expt7: Redundancy in REG

Describe a target person,
N=155, 28 Waldo images

How many landmarks are
mentioned, given visual
properties of target and
scene?



Woman with blonde hair in a pink top and red skirt in front of the man with papers coming out of his briefcase.

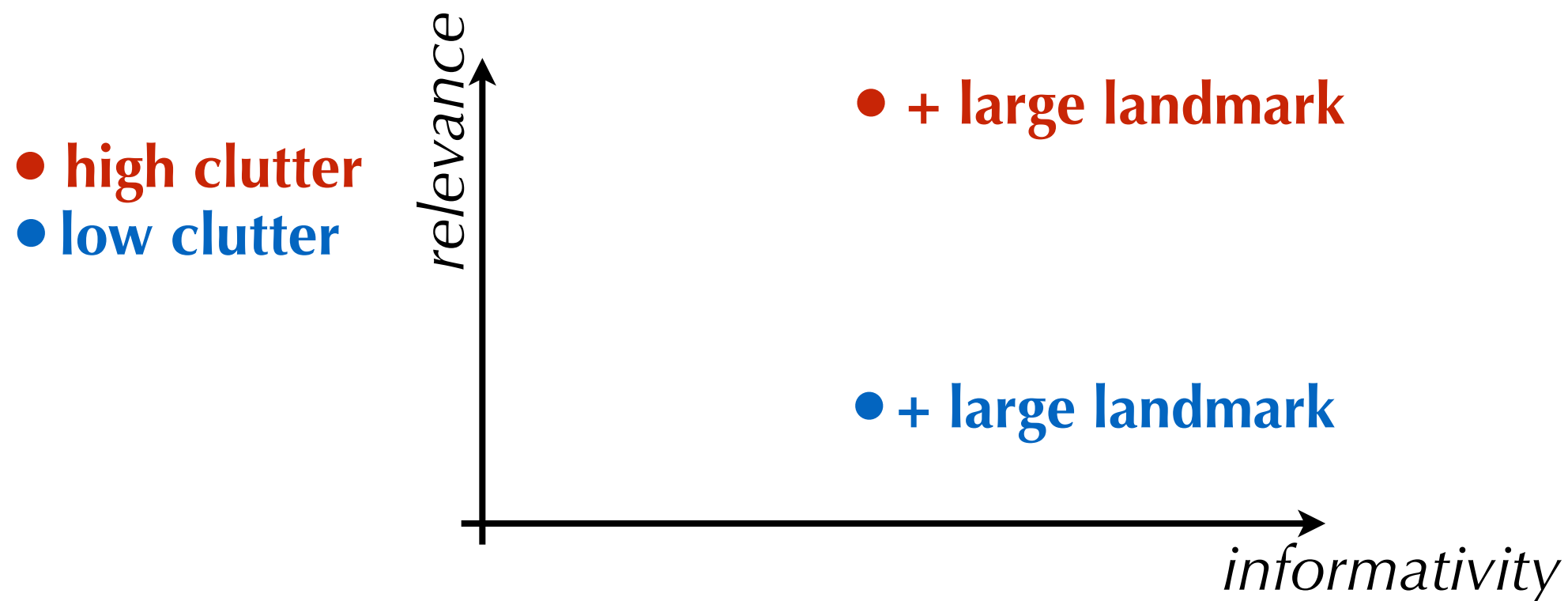
There is a man wearing a green jacket and red pants.

Find Waldo! Northwest of him is a man in a gray jacket and brown pants. He is to the right of a woman with a yellow shirt/blue top, and to the right of the girl with the red top.

Expt7: Redundancy in REG

- ▶ 85% of responses mentioned at least one landmark
- ▶ Targets with smaller area → more landmarks
- ▶ Targets with lower visual salience → more landmarks
- ▶ Scenes with more visual clutter → larger landmarks

→ Redundant landmarks are made relevant by visual scene and the task of visual search



Overspecification: The yellow pair?



- ▶ In **production**, speakers overspecify color more with clothing than with food
- ▶ Color may be more relevant to clothing (variable color) than food (constrained color)
- ▶ In **comprehension**, is the inclusion of a color adjective informative regarding object *category*?

Expt8: Redundancy in comprehension

Choose one of two pictures, N=19

Color Click on the yellow...

Control Click on the two...



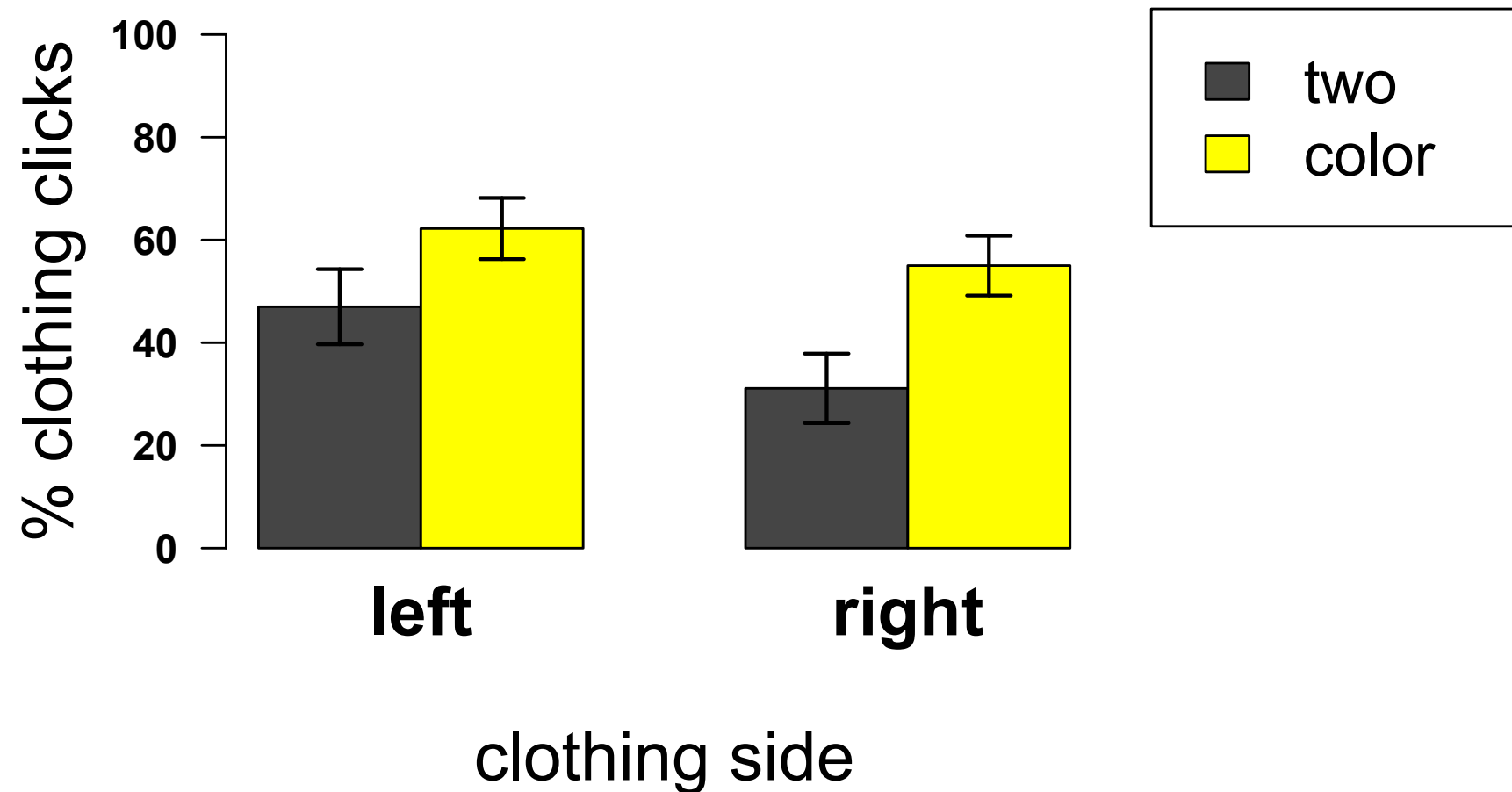
Balance left/right side of screen

- ▶ **Bigram frequency?**
“yellow shirts” vs “yellow bananas”
 - “yellow shirts” is more frequent, but so is “two shirts”
 - clothing is more frequent**Prediction:** clothing bias overall

- ▶ **Point-wise mutual information?**
yellow~bananas vs yellow~shirts
 $PMI(\text{red}, \text{cherries}) > PMI(\text{red}, \text{scarves})$
 $PMI(\text{purple}, \text{figs}) > PMI(\text{purple}, \text{heels})$
...
color~food > color~clothing
Prediction: color biases to food

- ▶ **Color as relevant to clothes**
Prediction: color biases to clothing

Expt8: Redundancy in comprehension



- Color is made relevant by properties of the object category
- Comprehenders are informed by “uninformative” color

Why are you (bothering) telling me this?

- ▶ **Relevance relations**
 - ▶ Cues to recovering relations
 - ▶ Repercussions for other phenomena
 - **Don't miss available relations or (machine-identifiable) cues**
- ▶ **Informativity**
 - ▶ Overly predictable messages
 - ▶ Useful redundancy in referring expression generation
 - **Not all redundant information is irrelevant**
- ▶ **Processing:** sweet spot for utterances that convey information that's
 - ▶ Plausible enough to be probable
 - ▶ Rare enough to be interesting
 - ▶ Relevant in context

► Thanks to:



Jet Hoek



Chris Cummins



Chris Lucas



Alasdair Clarke



Micha Elsner



Paula Rubio-Fernández

► And thank you!