

Production and comprehension of referential expressions show divergent behavior

Ming Xiang, Allison Kramer, Timothy Leffel and Chris Kennedy

The University of Chicago

An influential insight on language communication

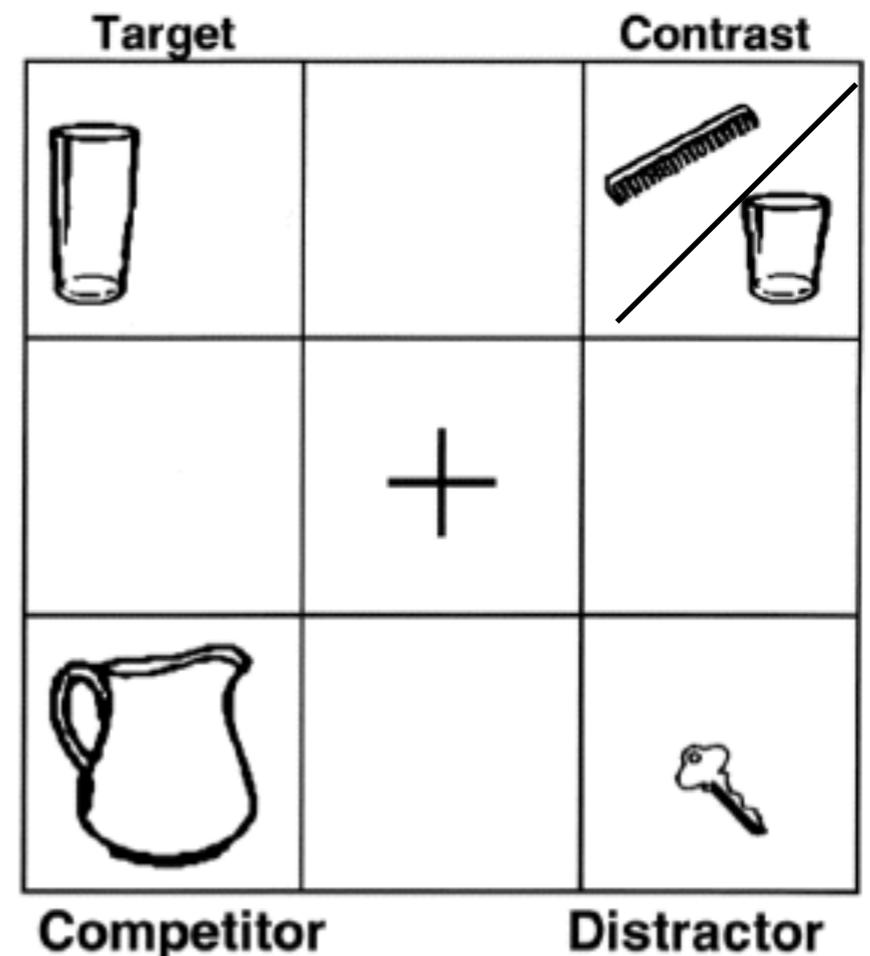
- Language communication involves coordination between speakers and hearers: speakers try to be informative, relevant and truthful; and hearers' comprehension of an utterance involves considering what could have been said by the speaker.
- These insights were informally presented by Grice, and were recently formalized using Bayesian framework (e.g. the Rational Speech Act model, Goodman and Frank, 2016)

$$P(\text{meaning}|\text{utterance}) \propto P(\text{utterance}|\text{meaning}) \times P(\text{meaning})$$

- A critical assumption behind this line of thought: There should be a close pragmatic alignment between production and comprehension behavior
- There has been relatively little empirical work exploring such alignment between production behavior and online comprehension
- The view that a “pragmatic” speaker chooses their utterances based on considerations about the hearer faces some empirical challenges that show production is egocentric and subject to production-internal pressure (Ferreira, 2008 for a review; Keysar et al., 2003; Lane et al., 2006; Engelhardt et al., 2006)

The referential contrast effect

- ▶ Sedivy et al. (1999) show that eye fixations converge on the target faster in the presence of a “contrast” object.
- ▶ Pragmatic speakers: say no more than is required for the purpose of communication, using as simple a form as possible (Horn, 1984; Grice, 1975).
- ▶ The eye tracking data reflects the pragmatic hearer’s belief about the referent based on their reasoning of a pragmatic speaker.



“Pick up the tall glass.”

Plan today

- Empirically evaluate the relationship between production and online comprehension of contrast effect
- Examine gradable adjectives with different scale properties (Rotstein and Winter, 2004; Kennedy and McNally, 2005; Kennedy, 2007; Burnett, 2012; Lassiter and Goodman, 2017)
- Manipulate the ordering between the target and the competitor objects in the display
- Collect data on both production and eye tracking

Preview of the findings

- There isn't a strong alignment between the contrast effect observed in production and in online comprehension (eyetracking).
- These results challenge the previous assumption that the contrast effect is best explained in terms of how a “pragmatic” hearer reasons about a “pragmatic” speaker.
- The findings motivate at least some domain specific processing mechanisms for the production and comprehension of the contrast effect

Testing the contrast effect on different adjective classes

Relative adjective: The threshold is always fixed in context. The evaluation of a relative adjective needs a contextually supplied comparison class.

“tall, short, big, wide, narrow, long, thick, thin”



“big boat”



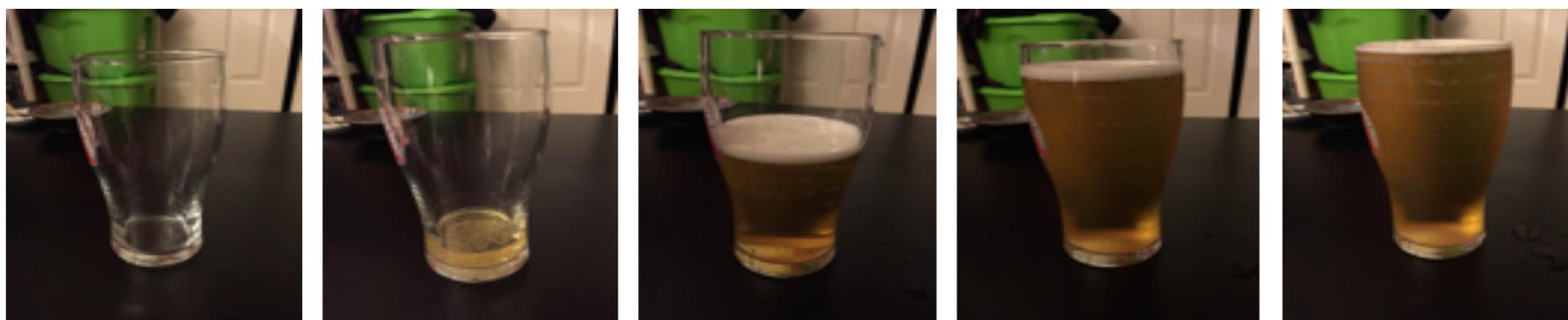
“tall candle”

Maximum adjective: The threshold defaults to the upper end of the scale, though contextual relaxation is possible.

“closed, empty, full, plain, smooth, straight”



“closed gate”



“full glass”

Minimum adjective: The threshold defaults to the lower end of the scale.

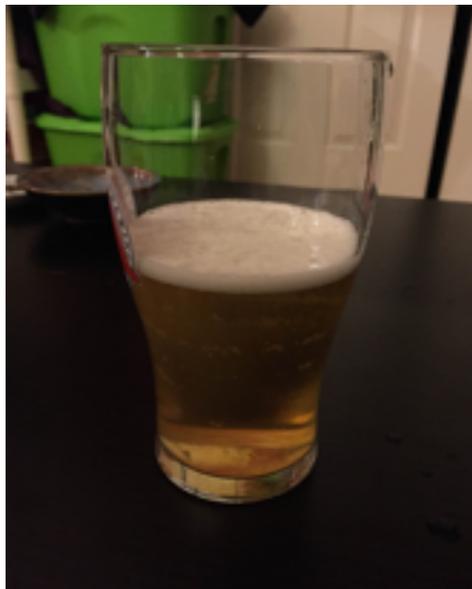
“bent, bumpy, curved, open, spotted, striped”



“bent nail”

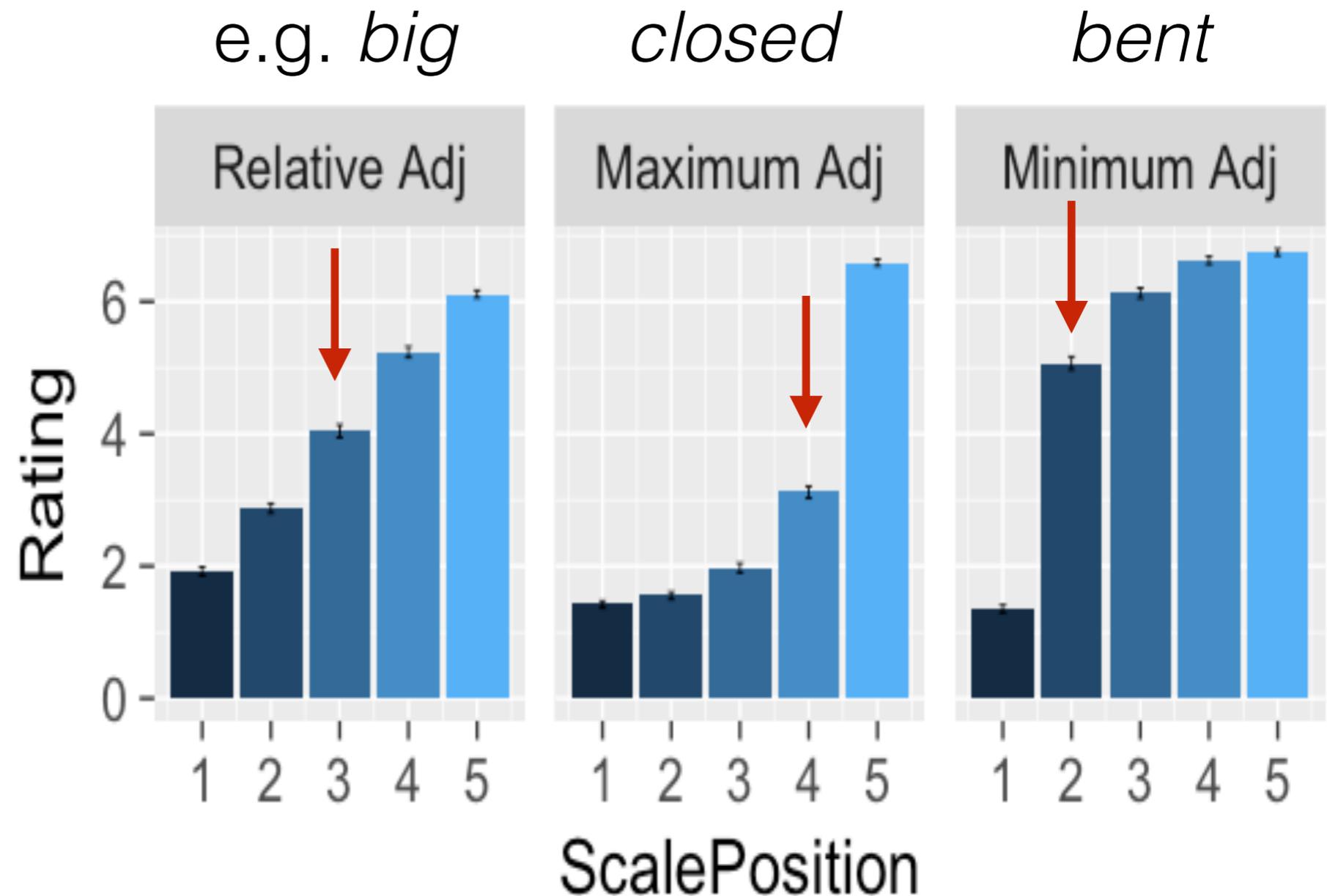


“spotted ladybug”



“This is a full glass”

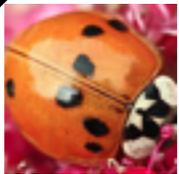
Norming rating: participants gave a 1-7 rating for each image on each scale we created.



The visual world eye tracking procedure

Stimuli example **T=C**

“Click on the **bent** nail”



Minimum Adj

“Click on the **big** boat”



Relative
Adj

“Click on the **closed** gate”



Maximum
Adj

Stimuli example

T < C

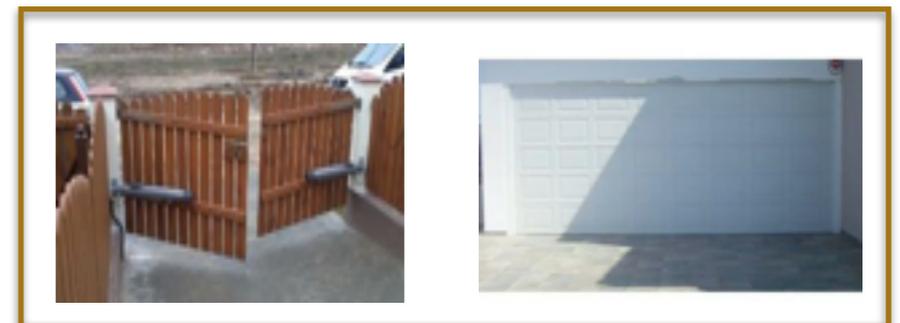
“Click on the **big** boat”

Relative
Adj

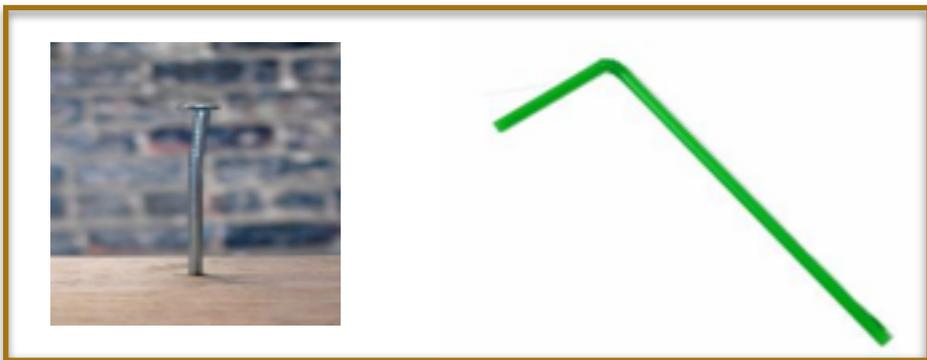


“Click on the **closed** gate”

Maximum
Adj



“Click on the **bent** nail”



Minimum Adj

Stimuli example

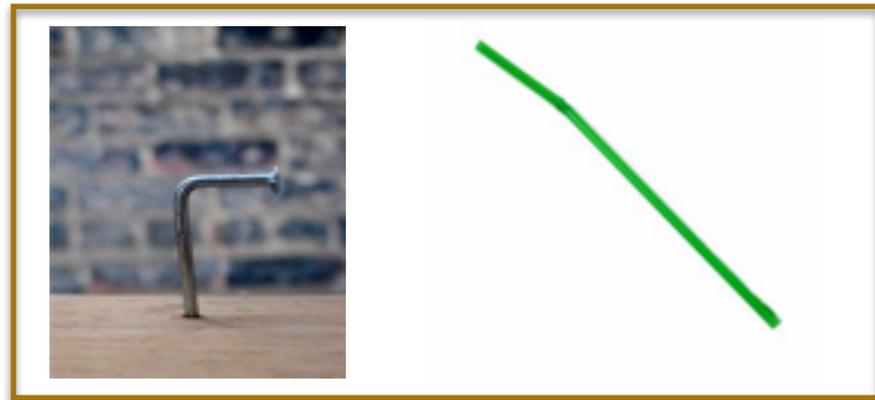


“Click on the **big** boat”

Relative
Adj



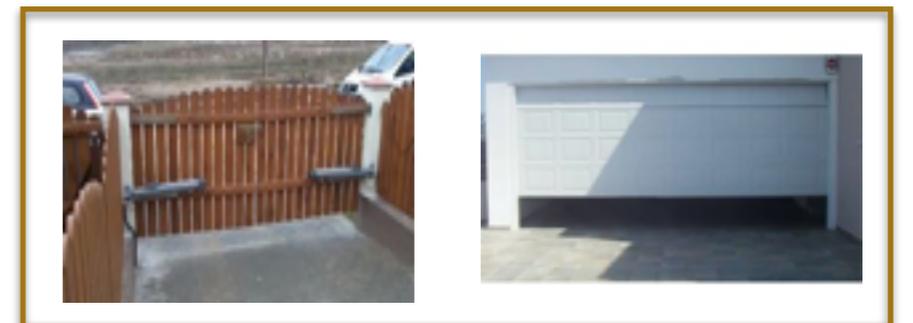
“Click on the **bent** nail”



Minimum Adj

“Click on the **closed** gate”

Maximum
Adj



Click on the **red** lipstick



In the filler trials, we also included trials to examine the contrast effect on color adjectives. But we did not manipulate the ordering relation between the target and the competitor.

The visual world eye tracking procedure

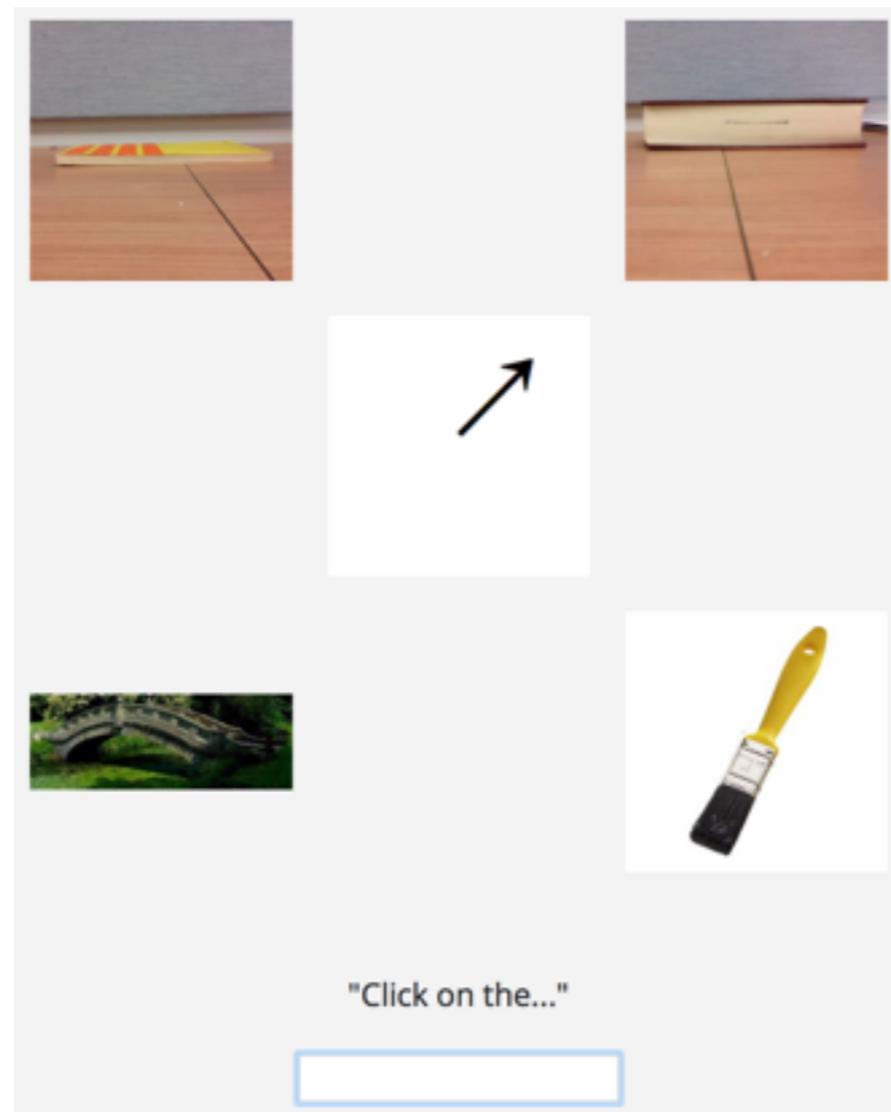
- Participants heard an instruction *Click on the adj. Noun*, and their eye movements were monitored.
- The $T=C$, $T<C$ and $T>C$ conditions were run in three separate experiments (with between 20-30 participants each)
- A female native English speaker produced all the audio instructions, and she was instructed to produce a neutral prosody as much as possible

Production procedure

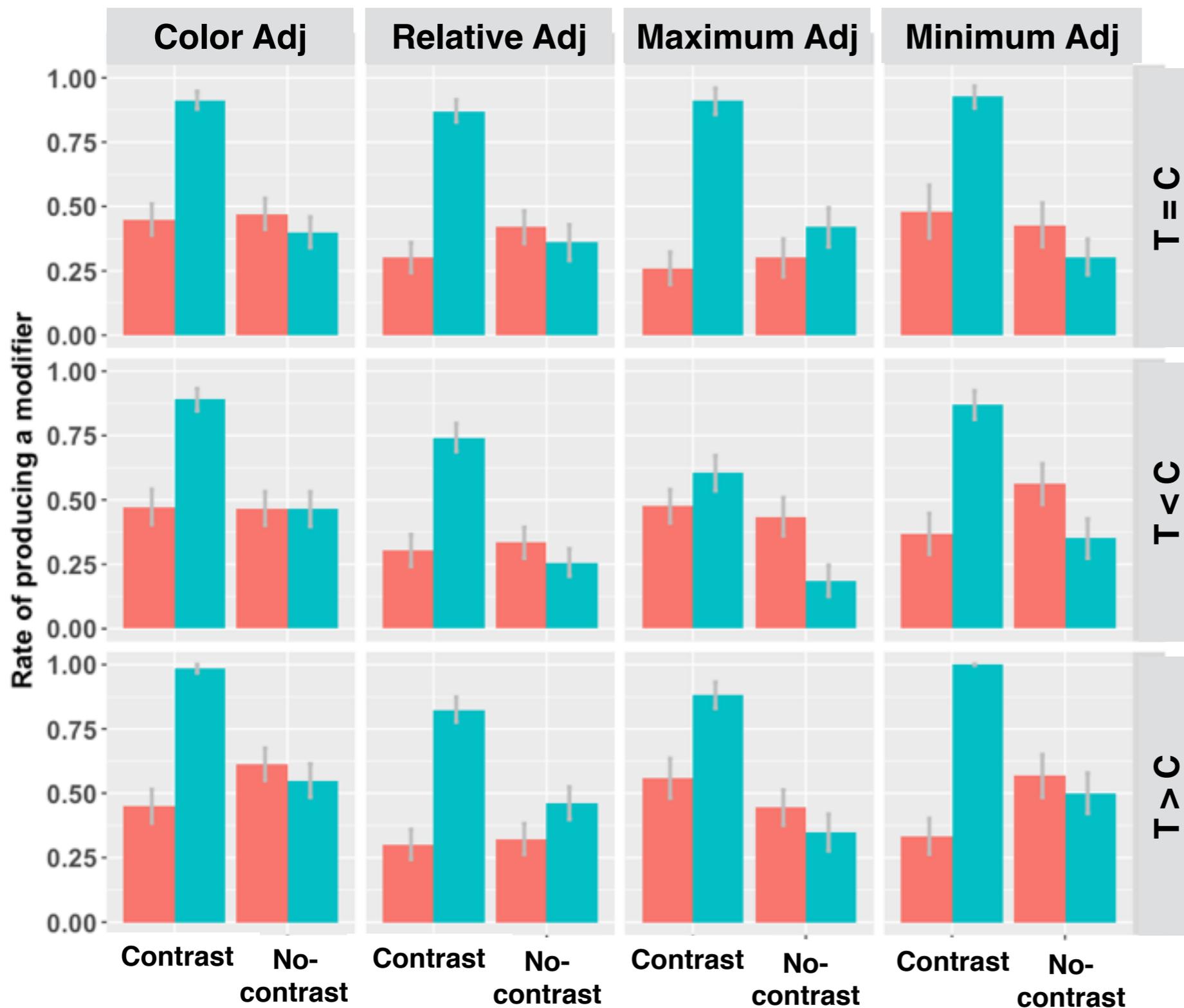
- Same set of visual stimuli as for the visual world comprehension study
- In each trial, participants were asked to describe an object in the visual display. Their responses were provided in writing. Data were collected on MTurk.
- The $T=C$, $T<C$ and $T>C$ conditions were run in three separate experiments (with 220 participants each)
- We analyzed the descriptions on the Target and the Competitor objects

An example trial for the production study

INSTRUCTIONS: Describe the object indicated by the arrow as if you are instructing a partner to click on it. Keep in mind that this partner can only see the images, not the arrow.



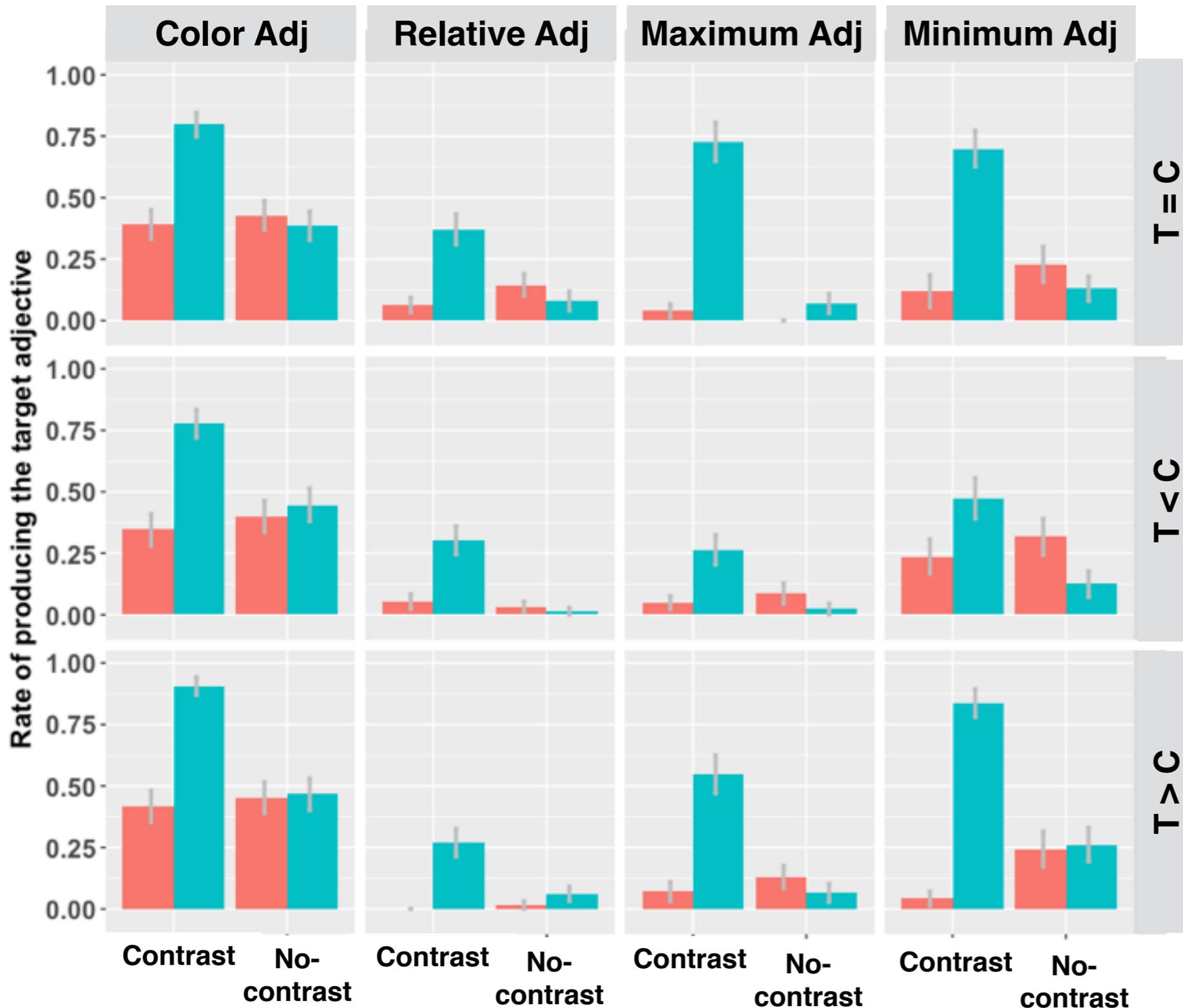
Rate of modifier production



■ competitor
■ target

Across the board, more modifiers were produced on the target in the contrast condition

Rate of key adjective production



■ competitor
■ target

- Across the board, more key adjectives were produced on the target in the contrast condition
- But, there are also clear effects of adj. type and ordering relation between the target and the competitor.

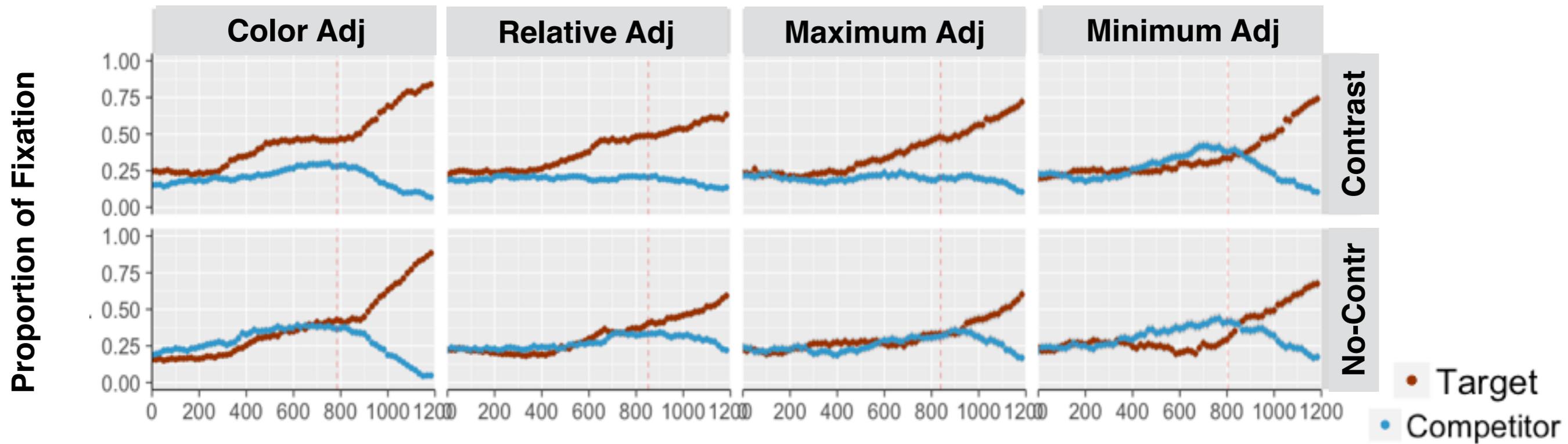
Summary of the production results

- Speakers were clearly “pragmatic”: more modifiers were produced on the target objects in the contrast condition, regardless of the degree relation between the target and the competitor objects
- Speakers were also clearly “pragmatic” when we just consider modifiers that are also the key adjectives of interest (i.e. the ones used in the comprehension study)
- However, there are also differences between different types of adjectives in terms of production frequency. We will come back to this observation later.

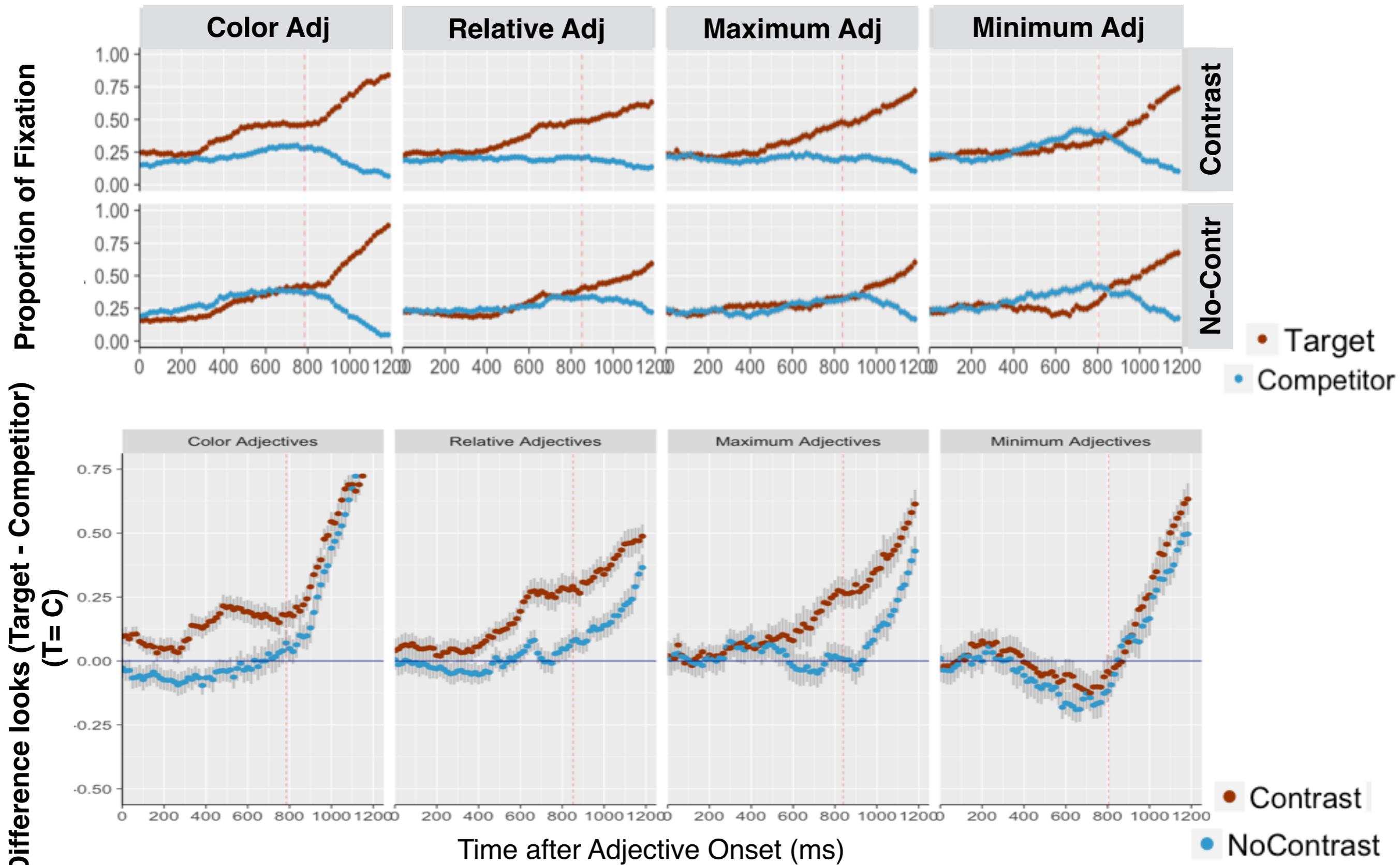
Expectation for a “pragmatic” hearer

- If there is a strong alignment between production and online comprehension of referential expressions, we should see contrast effect for all 4 types of adjectives (i.e. color, relative, maximum, and minimum adjectives), and all types of ordering relation between the target and the competitor.

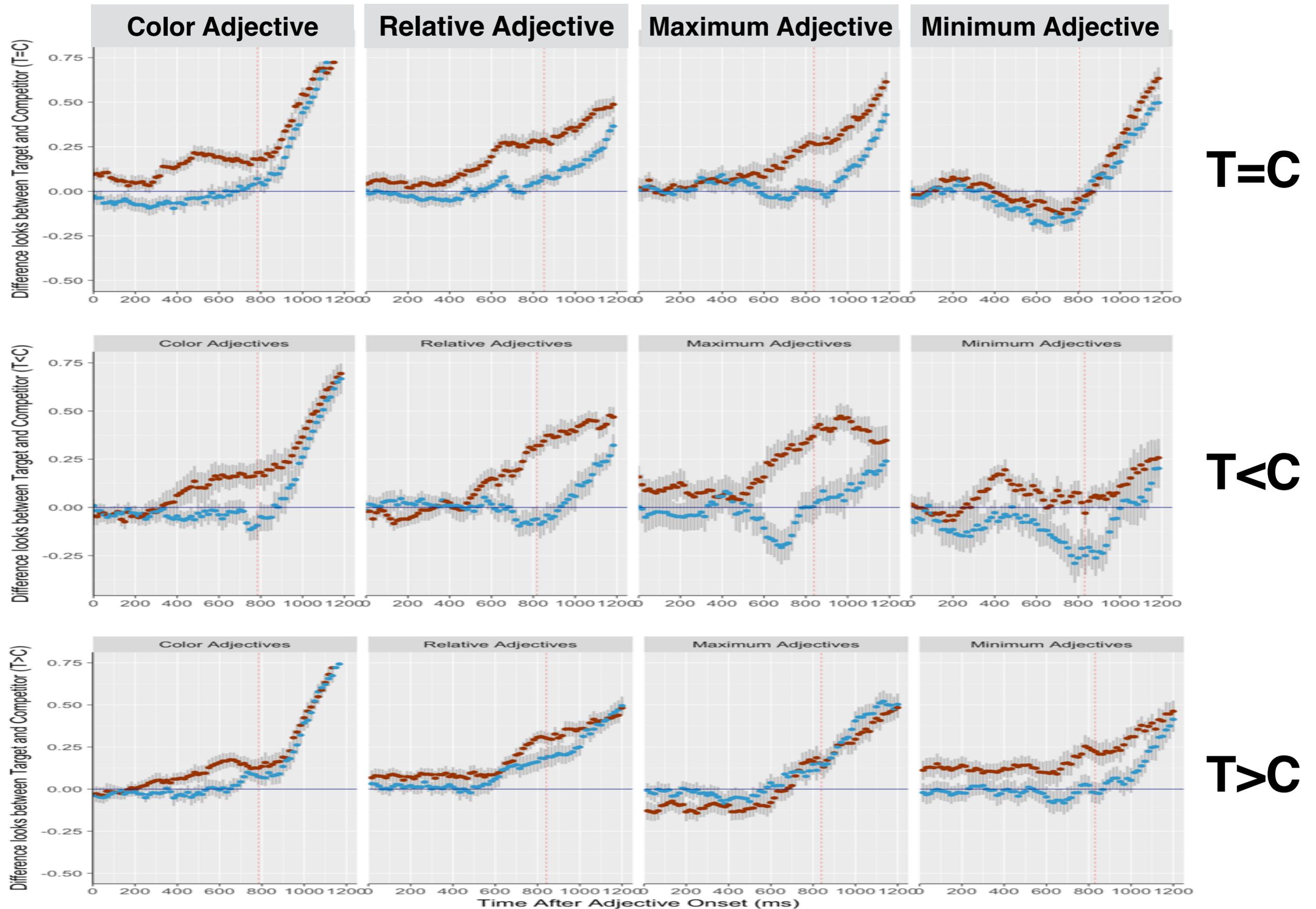
Eyetracking results (T=C)



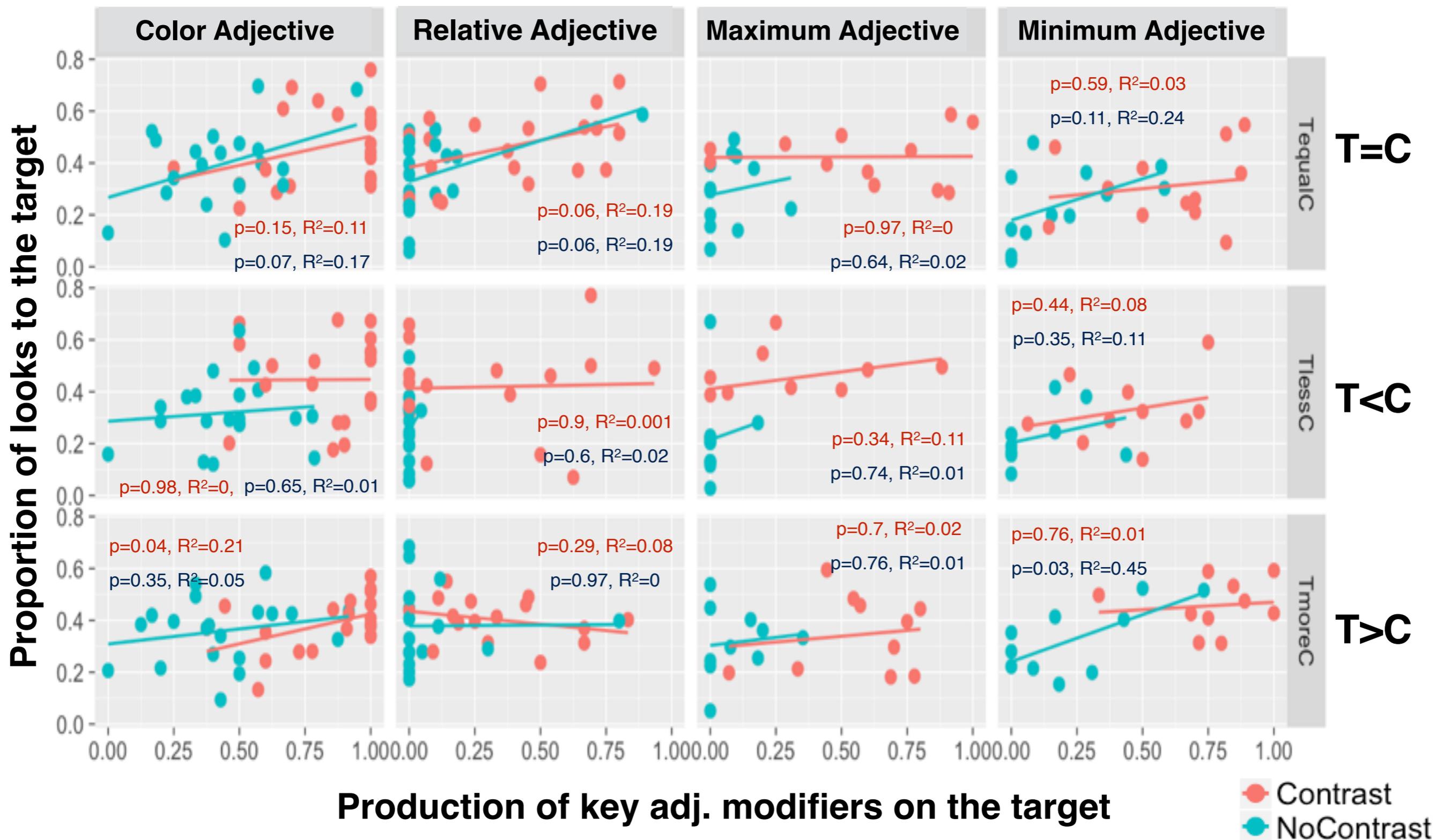
Eyetracking results (T=C)



Difference scores (Target - Competitor)



No robust by-item correlations between the production of key adjective modifiers and the proportion of looks to the target (the 200ms window before the noun onset was shown here; other time windows showed similar results):



Misalignment

- There is little alignment of modifier contrast effect in production and comprehension:
 - In production, speakers were strongly “pragmatic”
 - But the “pragmatic hearer” model doesn’t sufficiently account for the complete pattern of eye tracking data.
- Our eye tracking results are largely consistent with previous eye tracking results using geometric shapes (Leffel et al., 2016; Aparicio et al., 2015; Aparicio et al., 2017)

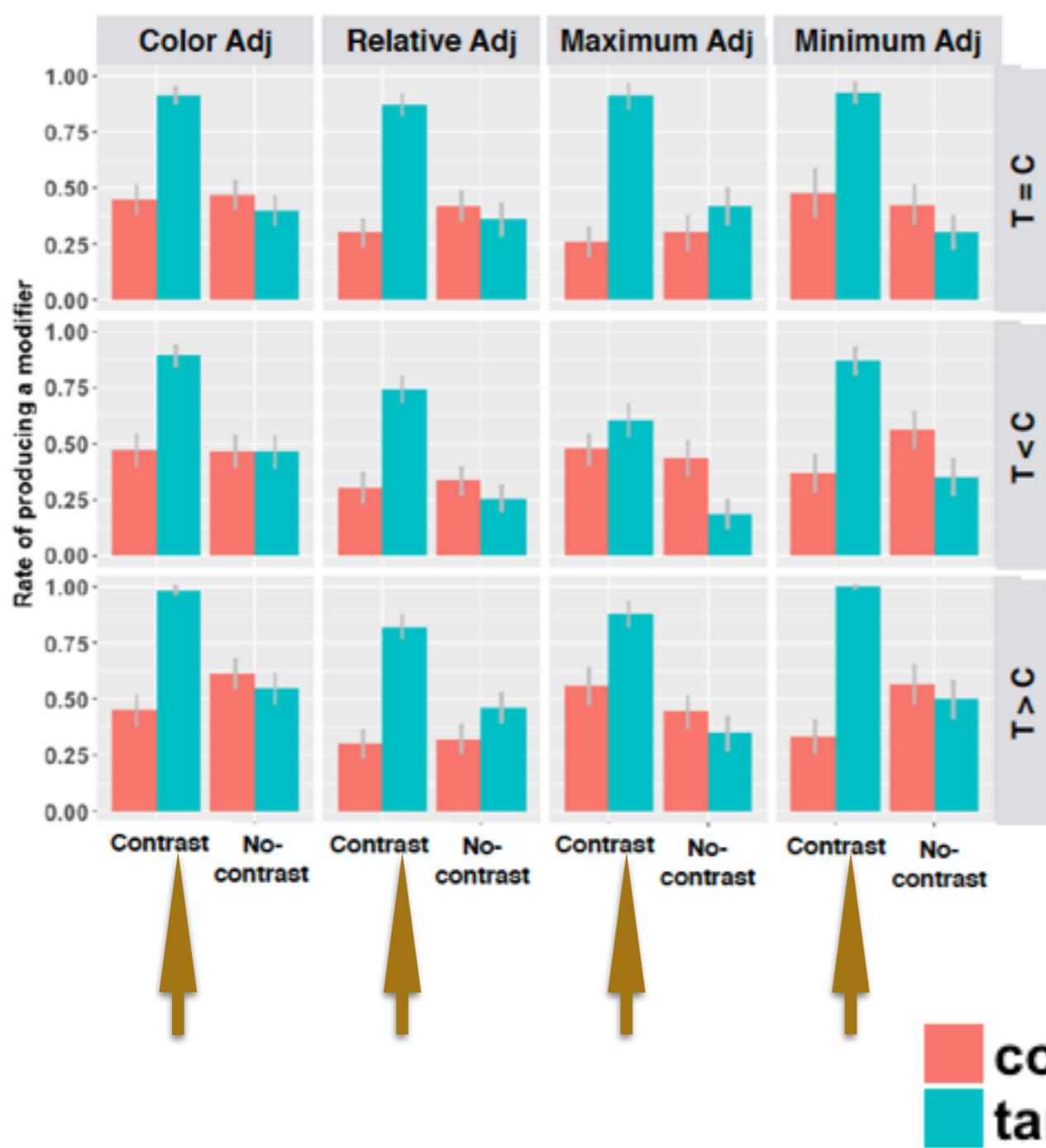
- The lack of strong correlation between production and online eye tracking data suggests a need to revisit the link between production and online eye tracking data.
- The findings challenge the idea that the eye tracking data reflects pragmatic hearers' belief about the referent based on their pragmatic reasoning of the speaker (see the poster on Fri. by **Qing, Lassiter and Degen** for a discussion on the similar question).
- We propose that there should be at least some semi-independent mechanisms to explain production and online comprehension (eyetracking).

Explaining the production results

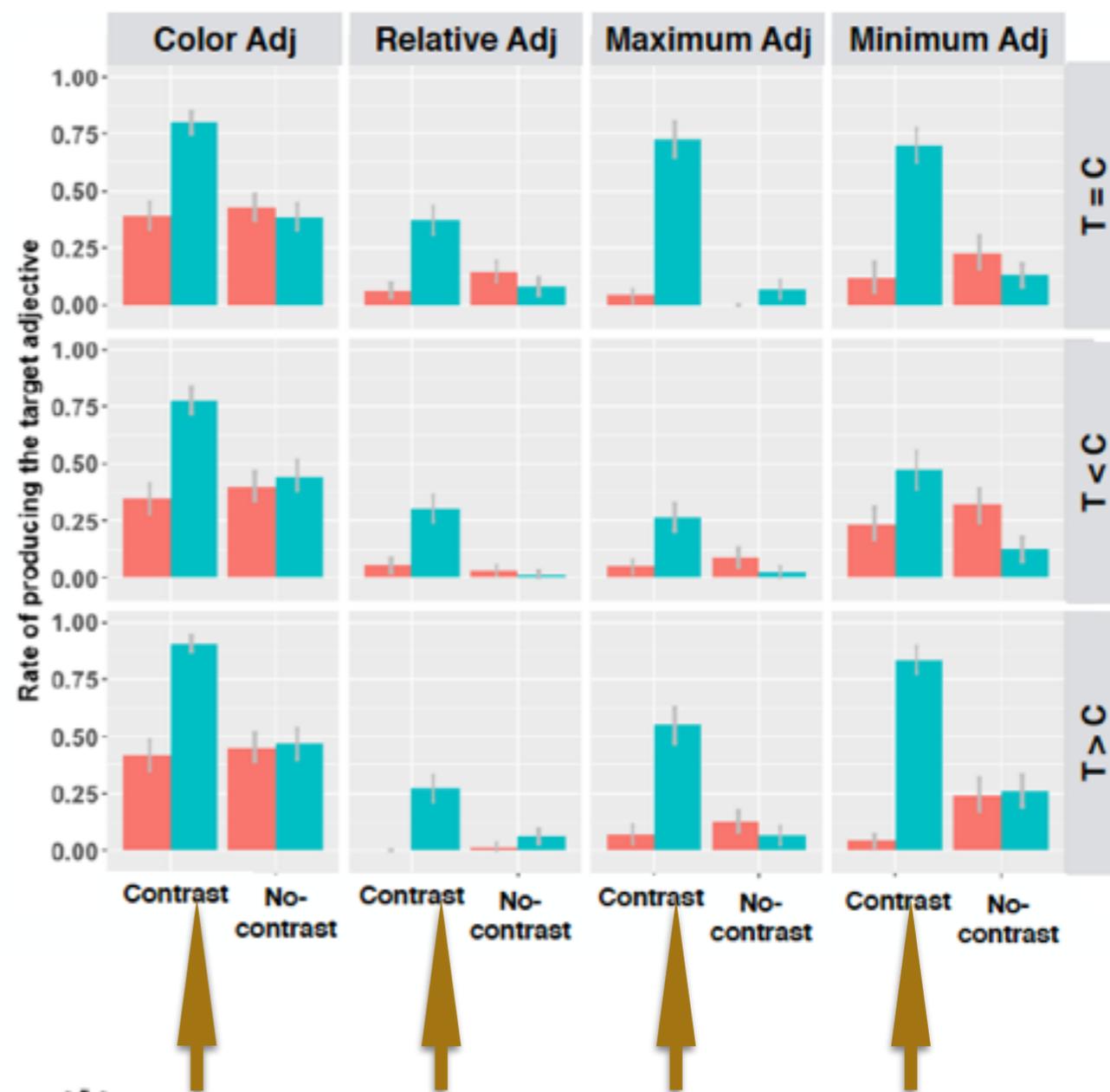
- We saw evidence that speakers are indeed “pragmatic”. But, the production system is also partly “accessibility” based (Bock, 1986; Ferreira and Dell, 2000).
- Loosely speaking, “accessibility” refers to how quickly a given lexical item can be successfully planned and retrieved and becomes available for the production system.
- We suggest that lexical semantic complexity is one of the factors that determine accessibility

What types of modifiers were produced on the target object in the contrast condition?

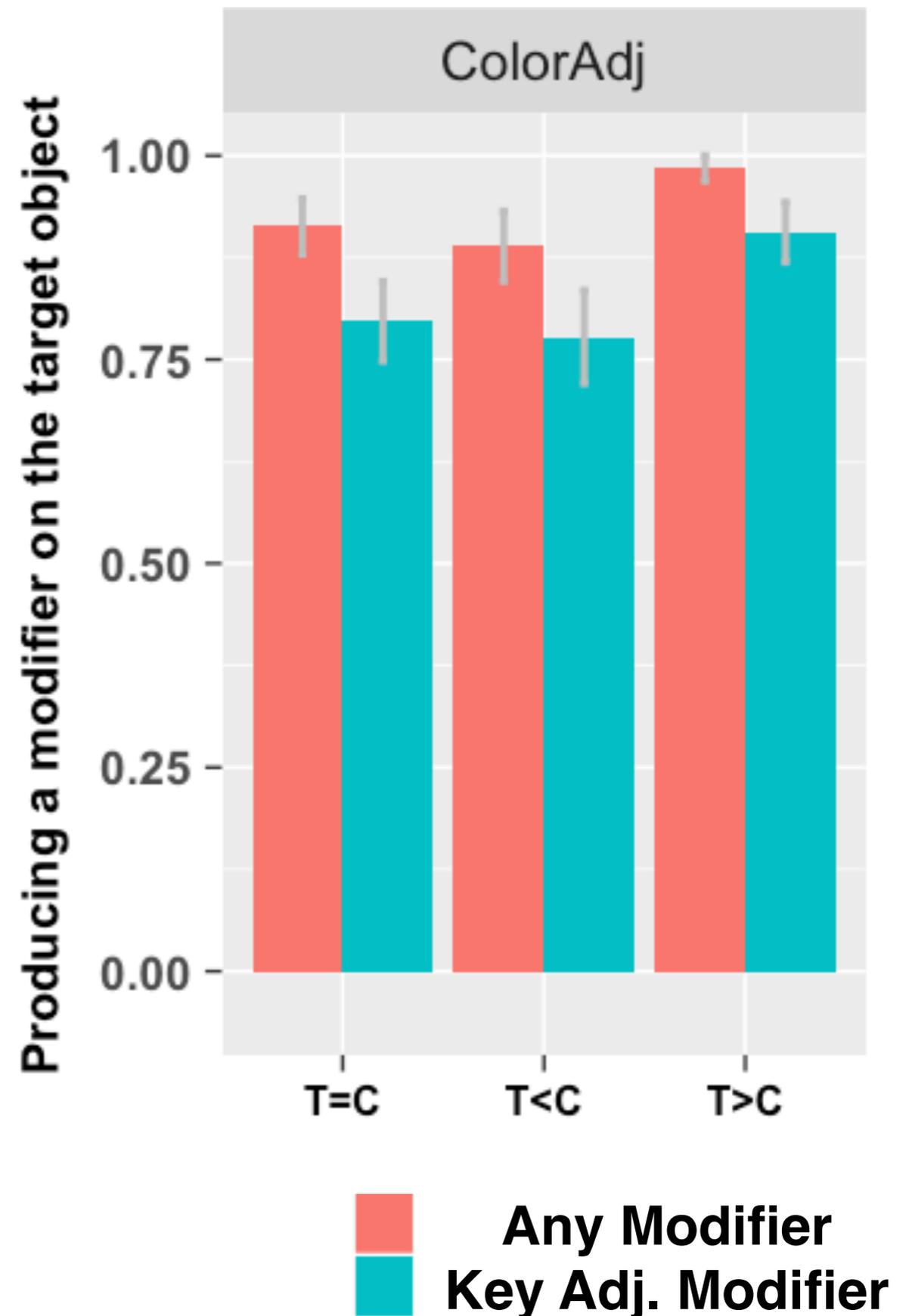
Rate of modifier production



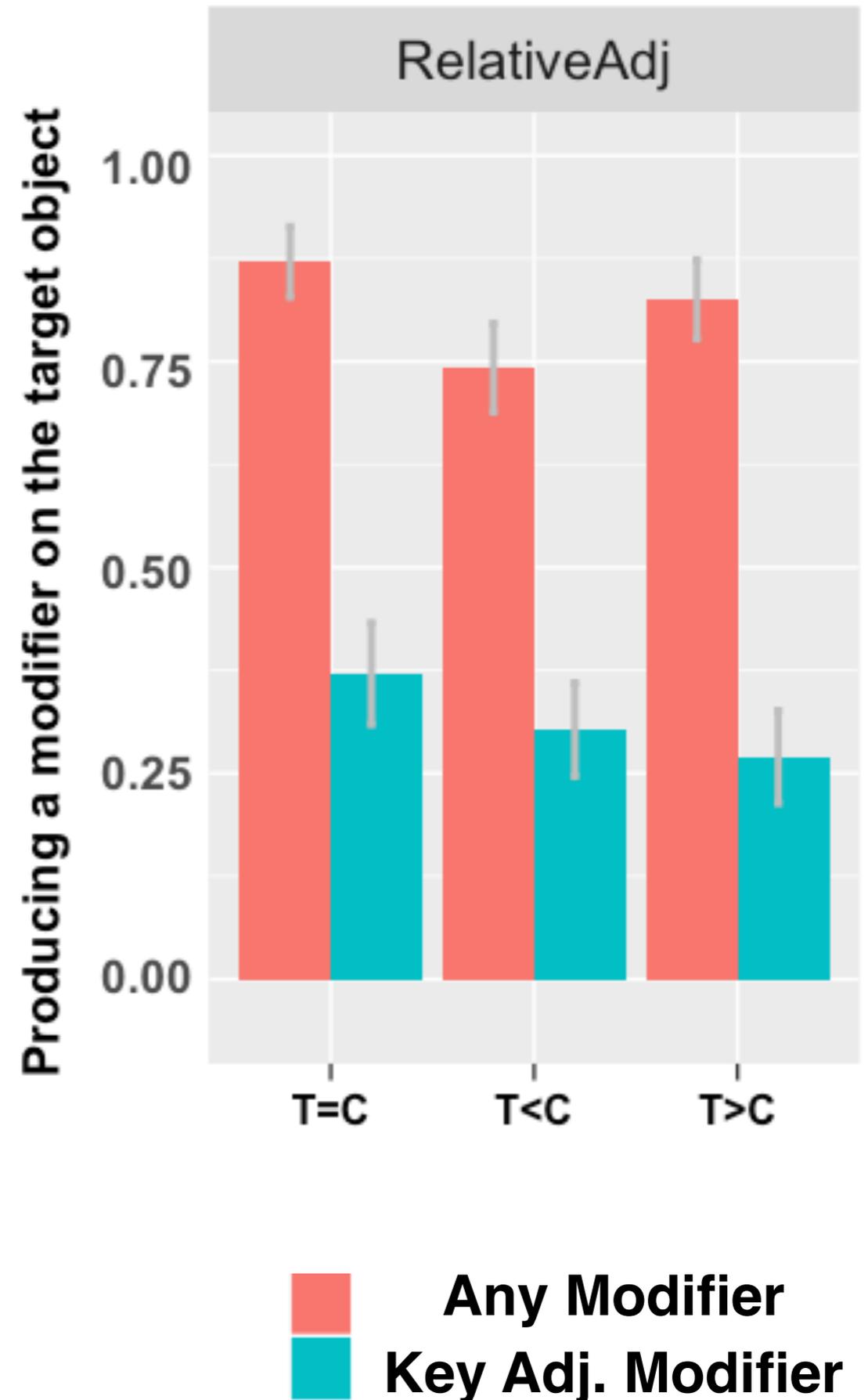
Rate of key adjective production



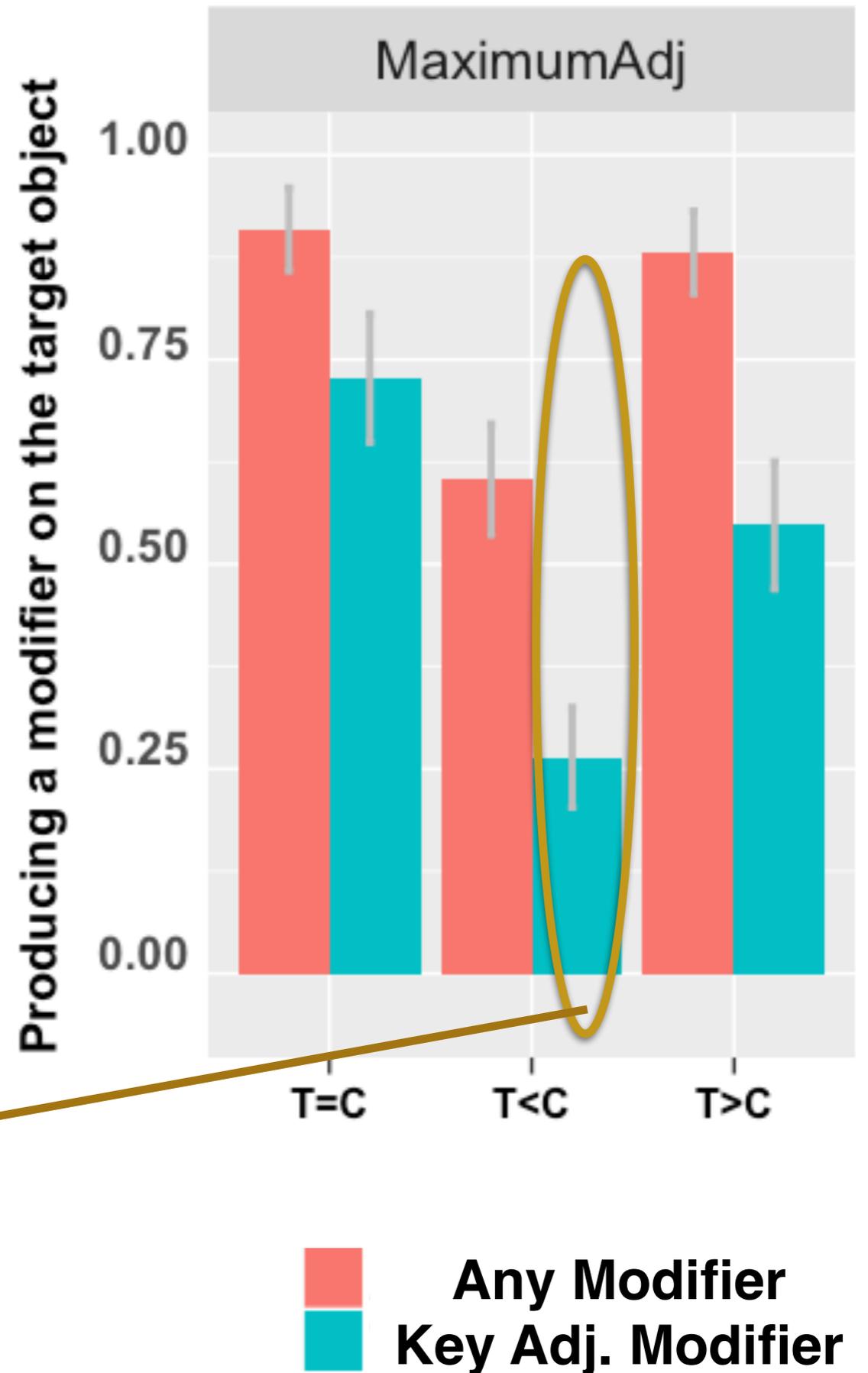
- **Color adjectives** (e.g. *red*) were the predominant choice of modifier when a modifier was produced in the contrast condition



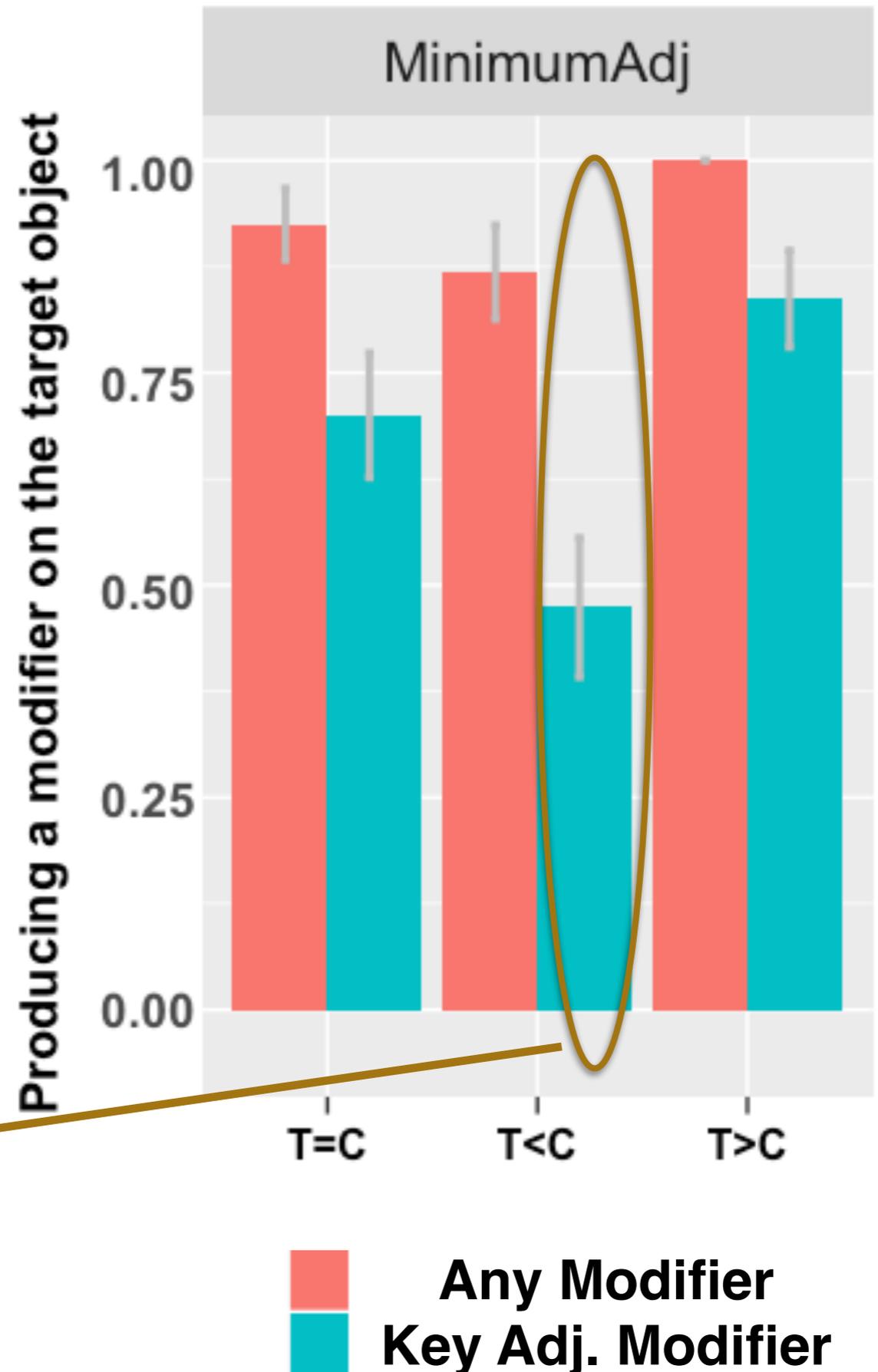
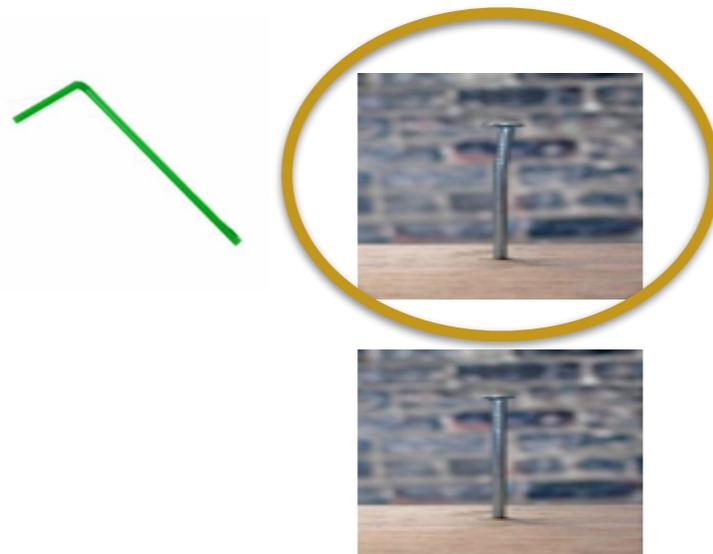
- **Relative adjectives** (e.g. *big*) were not the predominant choice of modifier in the contrast condition.
- Relative adjectives always require the computation of a flexible threshold, which enhances the complexity of the lexical retrieval process, and therefore may “lose” to other modifier candidates.



- **Maximum adjectives** (e.g. *closed*) were by and large the dominant choice of modifier when the target object is maximal.
- When the target object is non-maximal ($T < C$), speakers preferred to use a different modifier.
- Use of Max Adj. in $T < C$ would have required computation of flexible threshold as in Relative Adjectives.



- **Minimum adjectives** (e.g. *bent*) were the predominant choice of modifier when the target object was an optimal exemplar for the “Min” property.
- When the target object is not the most optimal ($T < C$), speakers preferred to use a different modifier.
- Use of Min Adj. in $T < C$ would have required computation of a context dependent threshold as in Relative Adjectives.



Conclusion for the production results

- Speakers are indeed pragmatic: they were rarely under-informative. More modifiers were produced when there is need to disambiguate between objects (i.e. in the contrast condition)
- But the specific kinds of modifiers speakers choose also reflect how quickly a given modifier can be planned and retrieved by the production system.
- When there is greater semantic and pragmatic complexity to determine whether a given modifier matches the target object, that modifier is less likely to be chosen

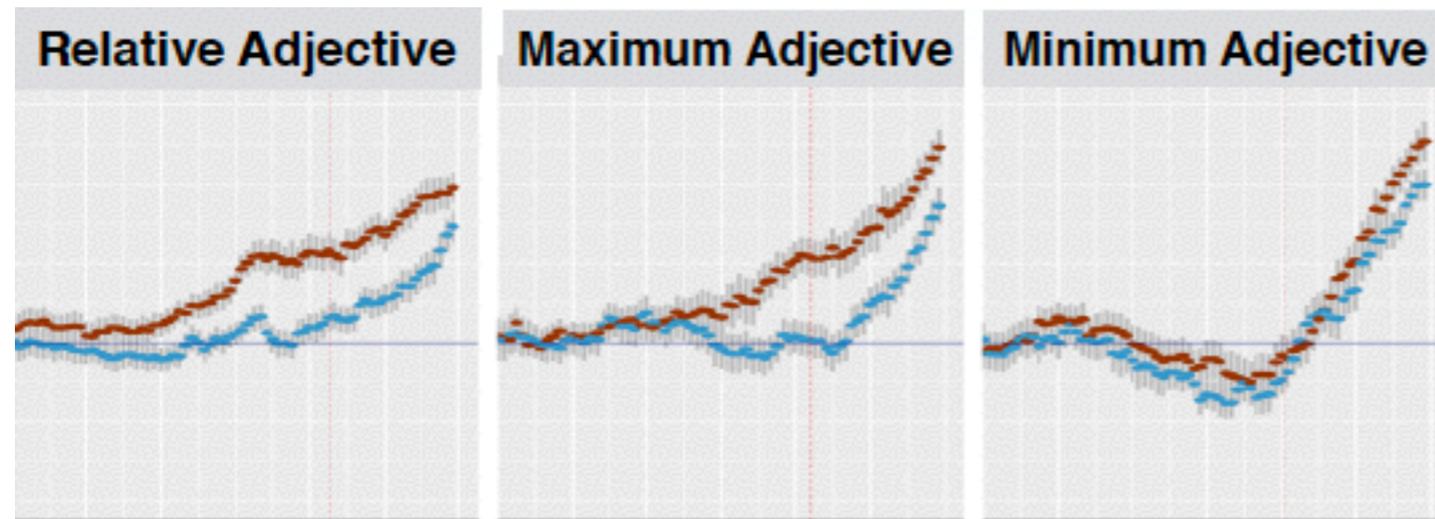
Some thoughts on the comprehension results

- The production and online comprehension results did not fully align.
- A complete account of the eye tracking results is still in development. We offer some preliminary thoughts here.
- The basic intuition is that online eye tracking data should reflect, at least partially, the interpretative process involved in deciding whether an adjective applies to an object.

- The interpretative process involves decisions about:
 - I. the adjective meaning (e.g. what counts as “*closed*” for a gate)
 - II. the relevant features of the objects (e.g. how closed a particular gate actually is)
- When there is uncertainty about either of these two decisions, the presence of a contrast object facilitates the computation. But when there is relatively high certainty, the contrast object doesn’t add additional benefit.

Let's take the **T=C** condition to illustrate the basic intuition.

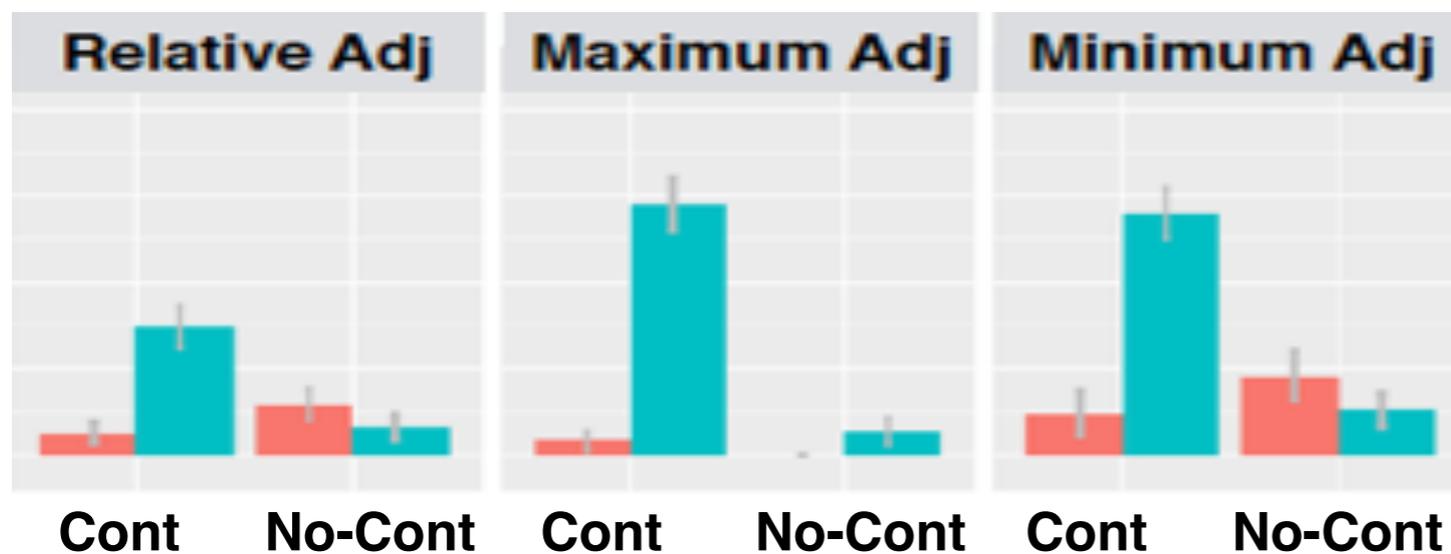
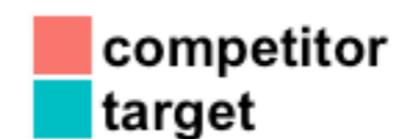
We can provide further discussion about the other conditions during Q&A.



Eye-tracking (Target - Competitor difference looks)



Production

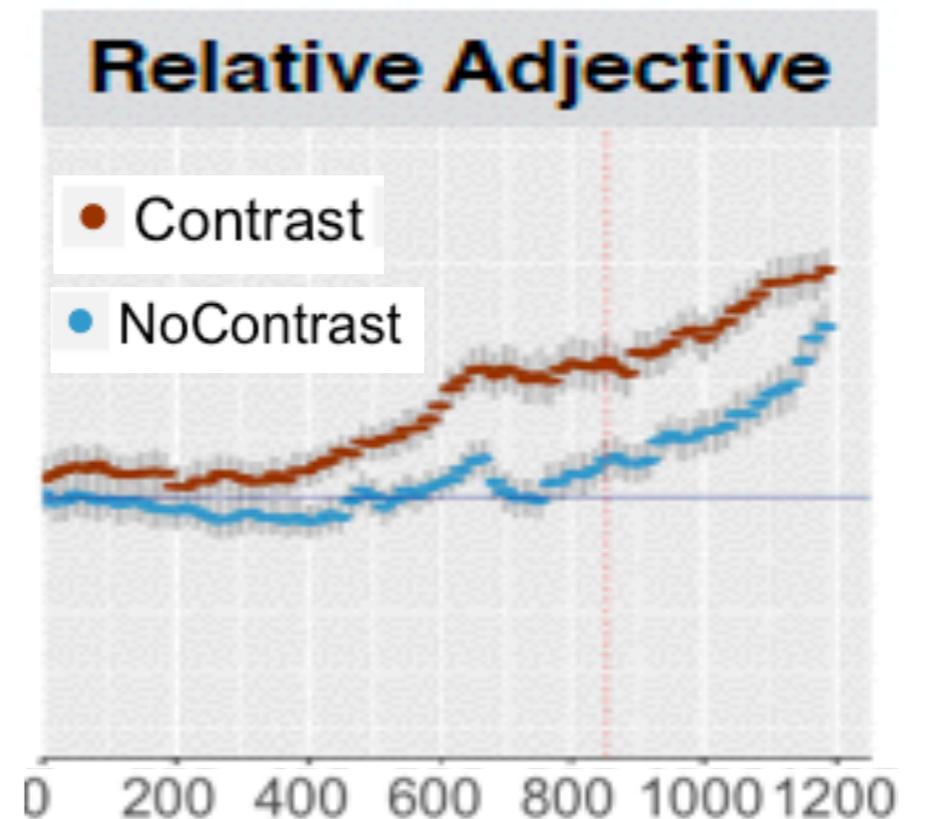


- Relative adjectives (e.g. *big*): $T=C$

When there is no contrast object, there is uncertainty about the adjective meaning (i.e. what counts as “big” for a given object), since the meaning of RAs are always computed relative to a comparison class.

The **Contrast condition** supplies a comparison class object explicitly to resolve the uncertainty about the adj. meaning

“Click on the **big** boat”

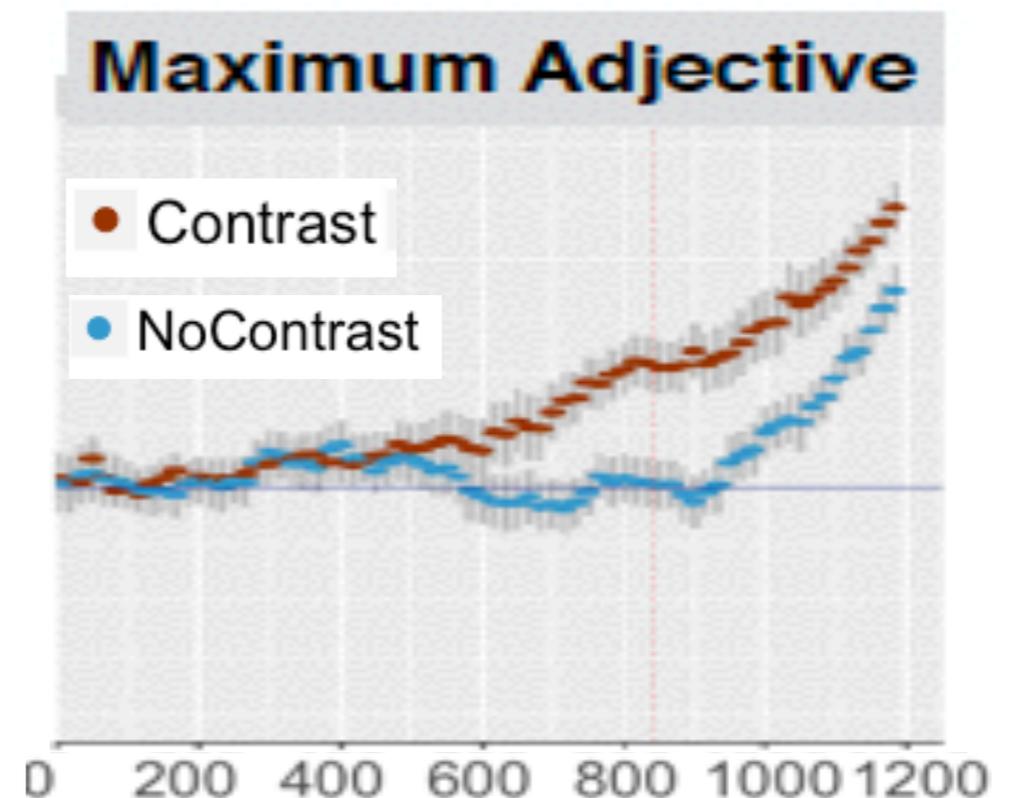


- **Maximum adjectives** (e.g. *closed*):
T=C

“Click on the **closed** gate”

When there is no contrast object, since strict satisfaction of the maximum adj. requires extreme precision, there may be uncertainty about the properties of the object (i.e. whether it has maximal degree). This triggers the related uncertainty about the adjective meaning (i.e. if a gate is not completely closed, to what extent “closed” could still apply).

The **Contrast condition** supplies a comparison class object explicitly to resolve the uncertainty in the No Contrast condition.

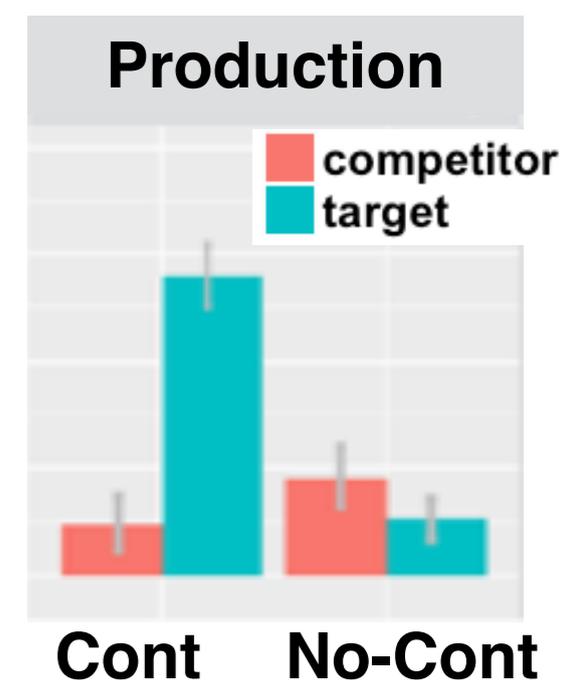
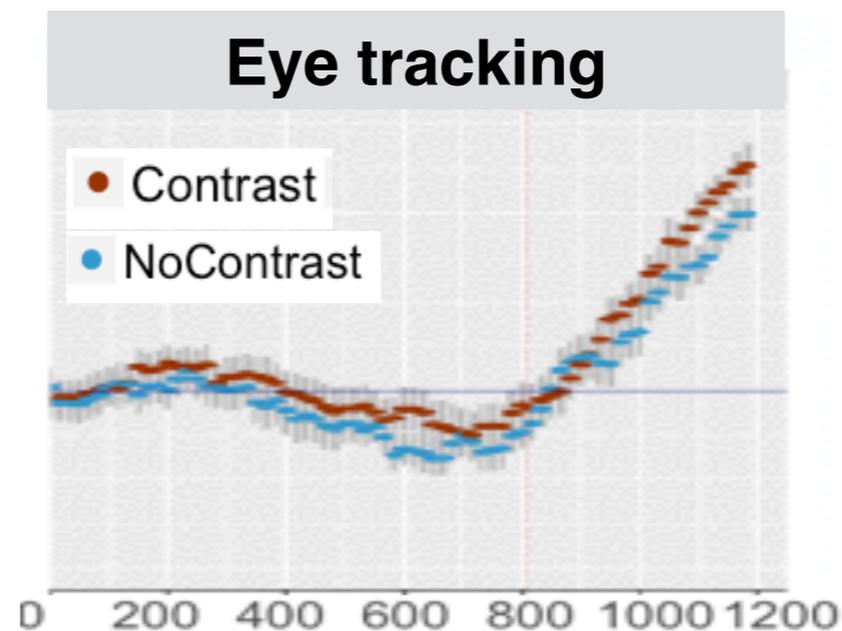
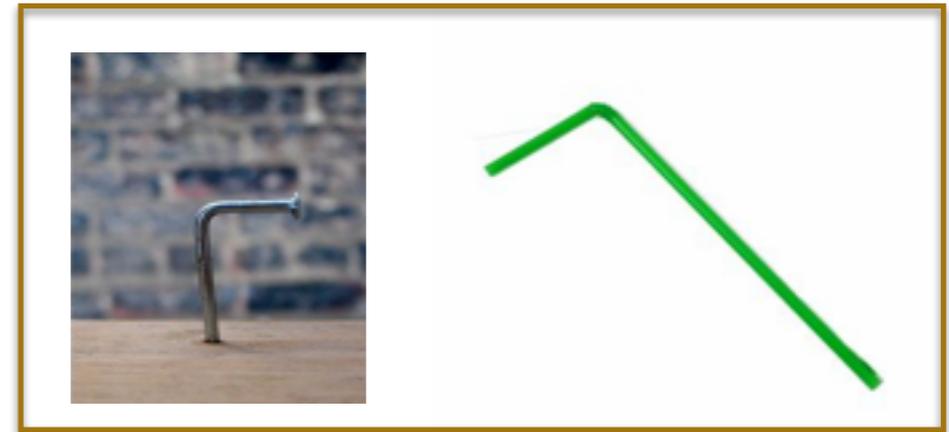


“Click on the **bent** nail”

- **Minimum adjectives** (e.g. *bent*): $T=C$

Even when there is no contrast object, there is already **high certainty** about both the meaning of the minimum adj (i.e. what counts as bent); and about the properties of the object (i.e. whether it has some bend).

There is no additive benefit from the contrast object.



Conclusions

- There isn't strong alignment between production and online comprehension. This raises questions for the basic architecture of language processing.
- We propose that the production system is constrained by the “accessibility” of to-be-produced material, even though speakers are also indeed “pragmatic” and are rarely under-informative.
- The eye tracking results for the contrast effect *does not* reflect pragmatic hearers' belief about the referent based on their reasoning of a pragmatic speaker. Instead, it reflects how cues from different sources (e.g. linguistic, visual context, etc) are combined to arrive at interpretive decisions.

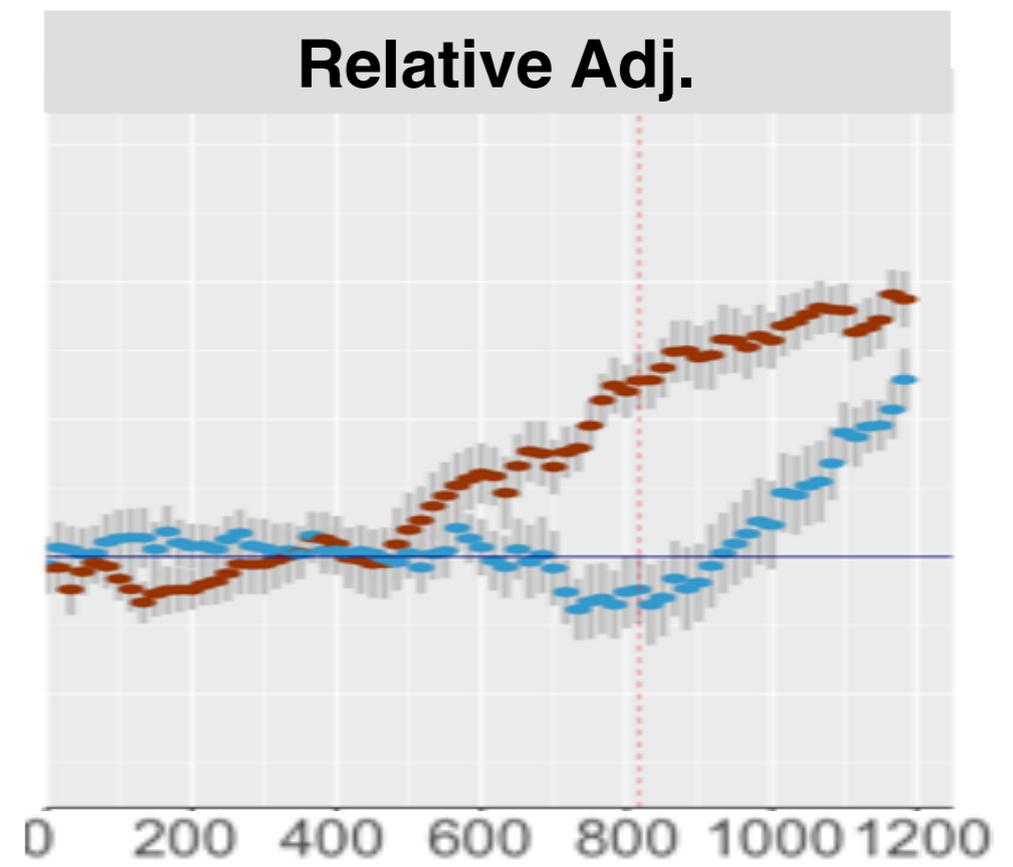
We want to thank the grant support from NSF BCS 1227144; Josef Klafka for eye tracking data collection and coding the production data; and Ciyang Qing, Judith Degen, Dan Lassiter for discussion.

Thank you!

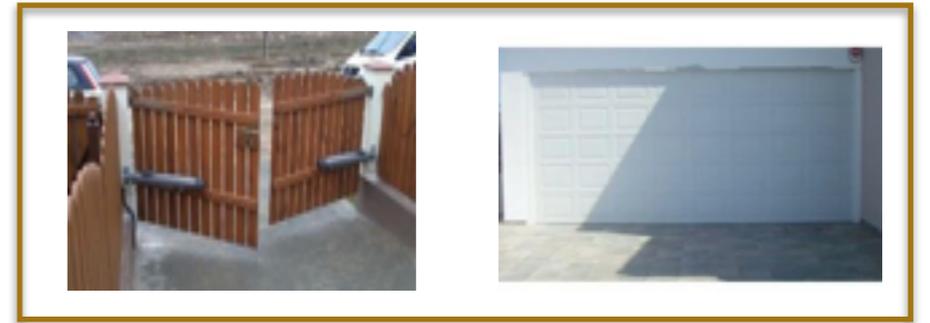
“Click on the **big** boat”



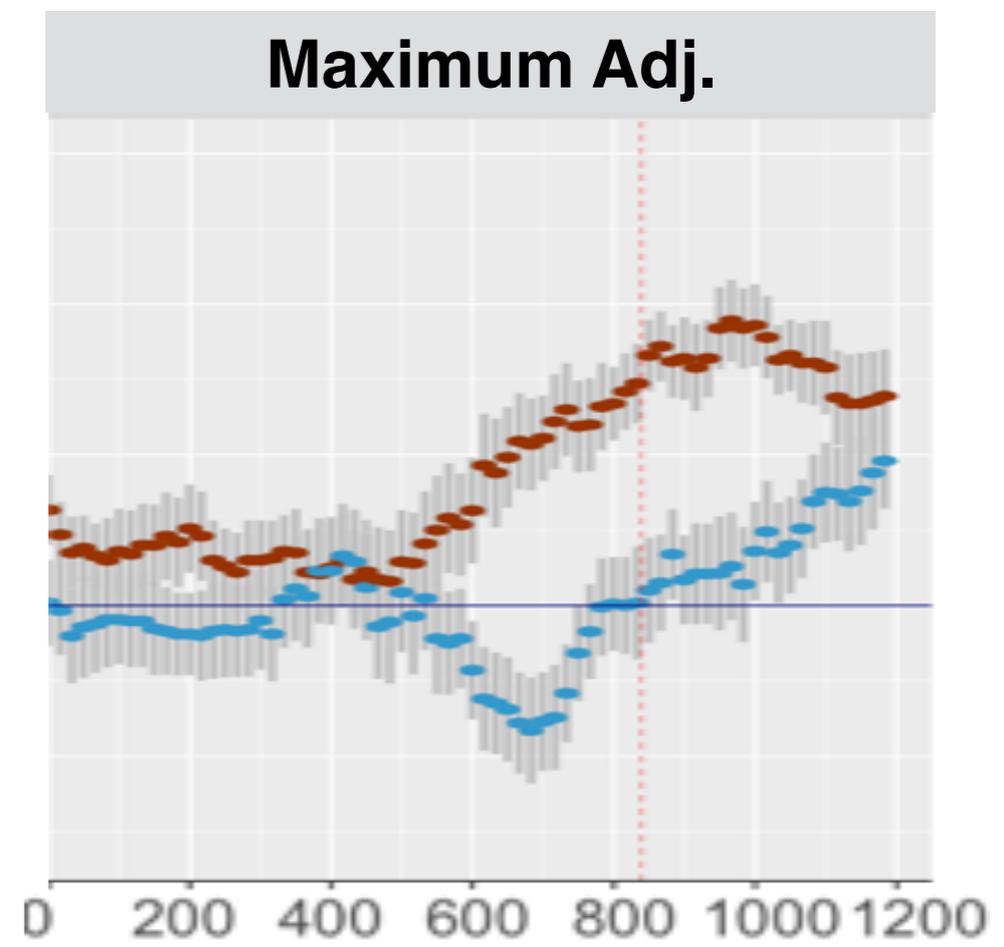
- Relative adjectives (e.g. *big*): $T < C$



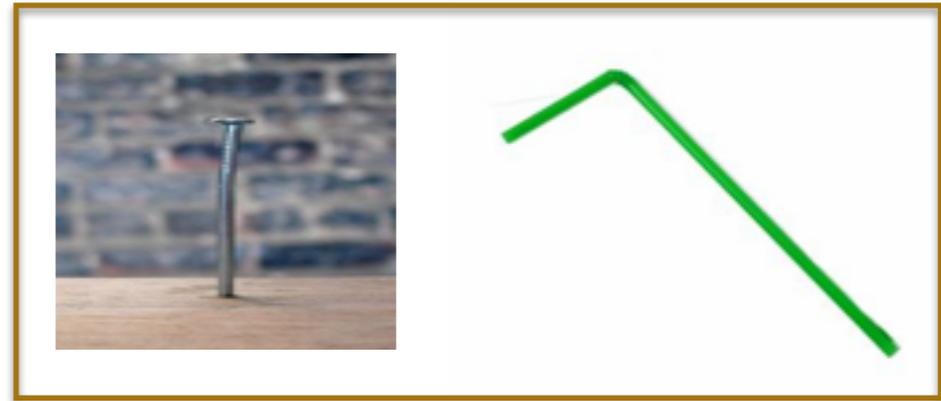
“Click on the **closed** gate”



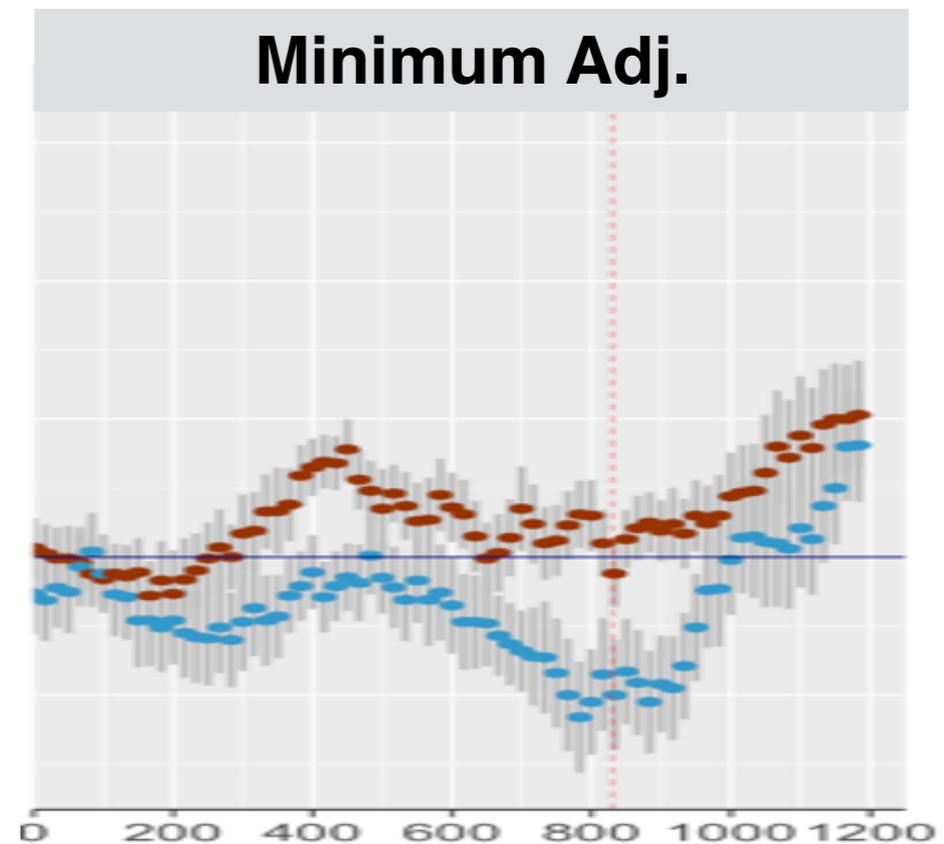
- Maximum adjectives (e.g. *full*): $T < C$



“Click on the **bent** nail”



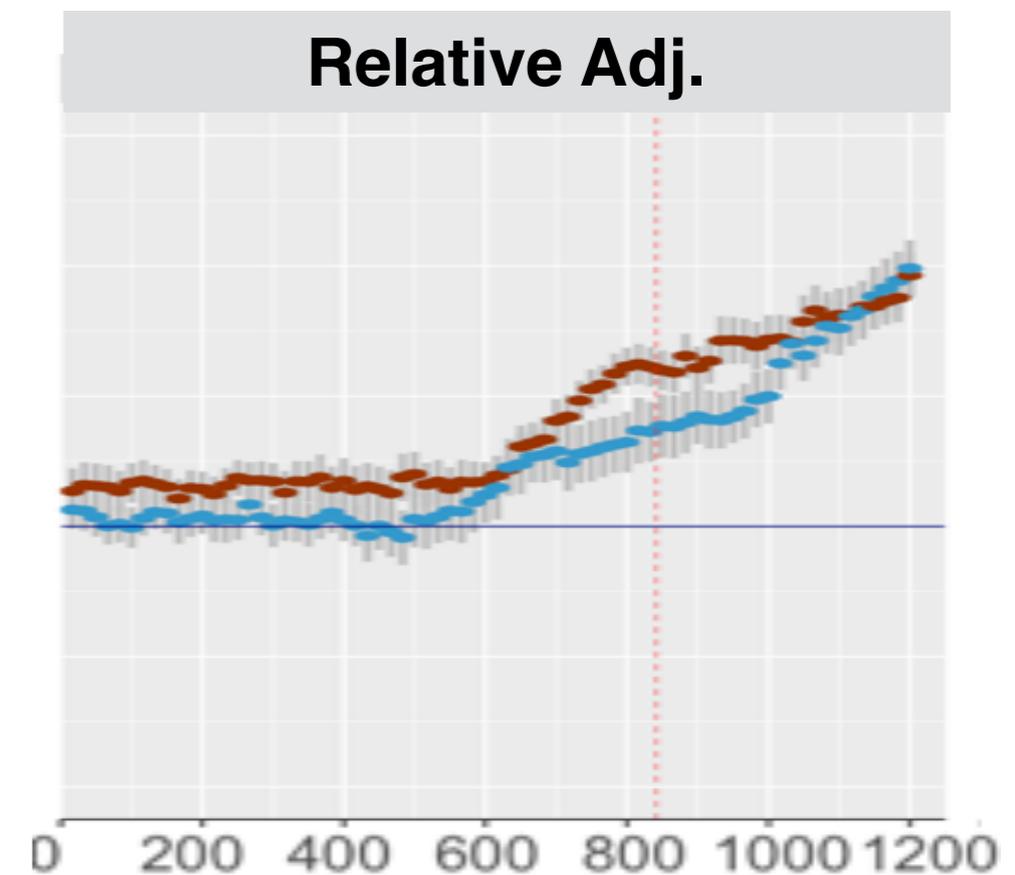
- **Minimum adjectives** (e.g. *bent*): $T < C$



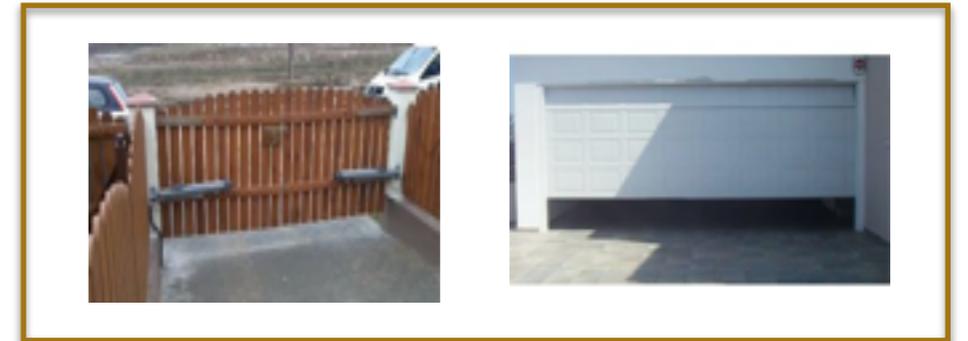
“Click on the **big** boat”



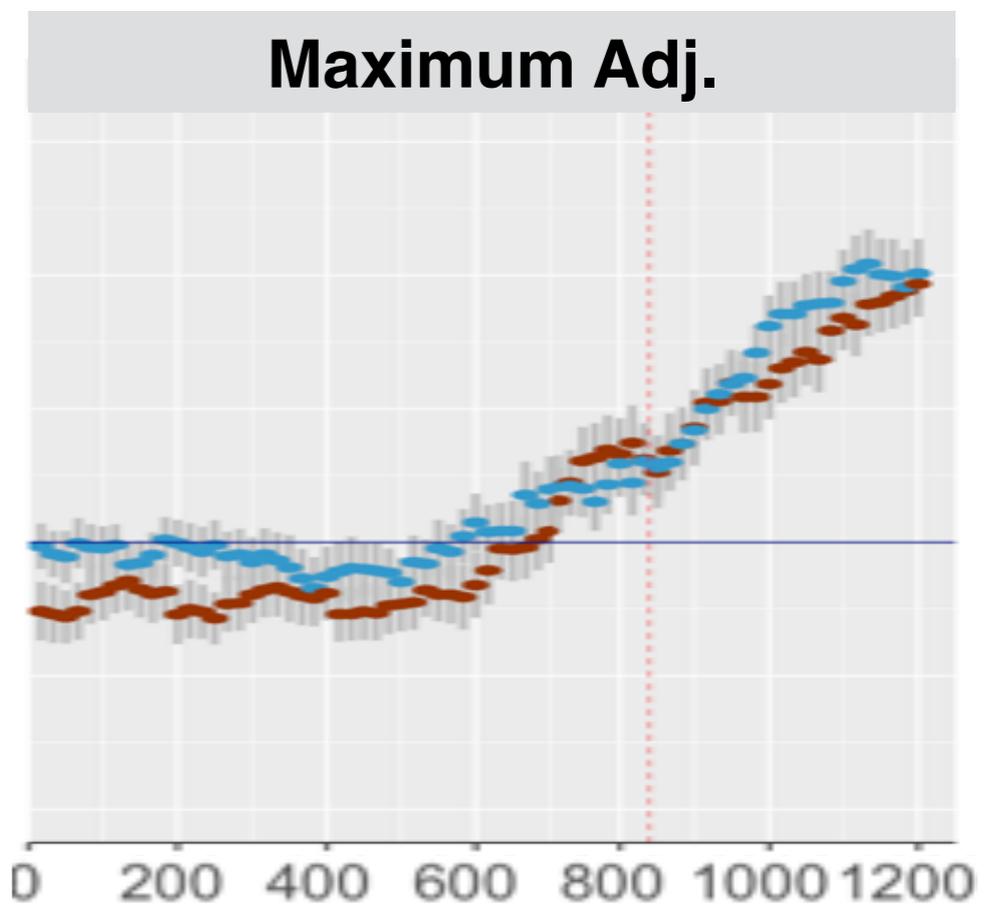
- Relative adjectives (e.g. *tall*): $T > C$



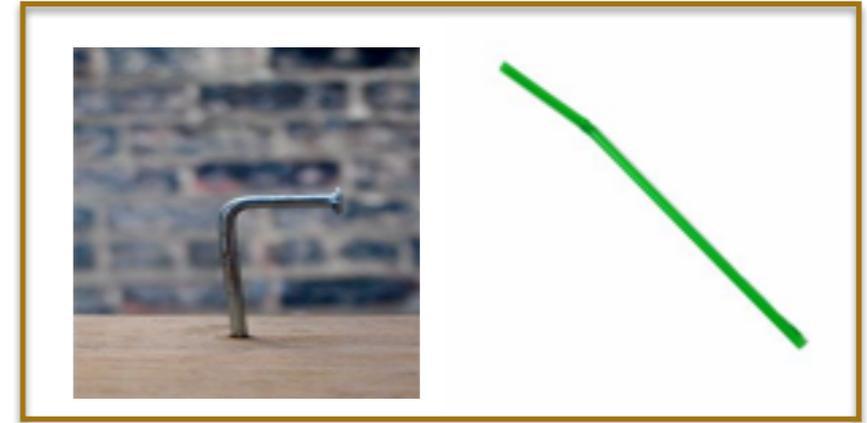
“Click on the **closed** gate”



- Maximum adjectives (e.g. *full*): $T > C$



“Click on the **bent** nail”



- **Minimum adjectives**
(e.g. *bent*): $T > C$

