Production Expectations Modulate Contrastive Inference

Anonymous CogSci submission

[ek: TODO: generally make axis labels bigger]

Abstract

[ek: TODO: needs to be shortened to 150 words!][jd: cut it down. need to cut 14 more words.] Contrastive inferences, whereby a listener pragmatically infers a speaker's referential intention by reasoning about other objects in the context that the speaker might be referring to with a partial referring expression like *the yellow*, are notoriously unstable. We report a production-centric model of interpretation couched within the Rational Speech Act framework. Adjective production probabilities a listener expects for objects in a given context drive the presence and size of c ontrastive inferences: the greater the asymmetry in expectation for a speaker to use a pre-nominal adjective for the target rather than for competitors, the greater the listener's resulting target preference. Modifier production probabilities collected in a free production study (Exp. 1) were used to make predictions about comprehension in an incremental decision task (Exp. 2). The model's fine-grained interpretation predictions are supported by the data. This account has the potential to explain the now-you-see-it-now-you-don't nature of contrastive inferences and shifts the explanatory focus away from contrastive inference narrowly and towards online interpretation of referring expressions more broadly.

Keywords: contrastive inference; RSA; typicality; incremental processing

Introduction

One of the most interesting features of language is its flexibility. In referring to an object, speakers choose from a wealth of possible referring expressions. *The banana*, *the yellow banana*, and *the curvy fruit* are all expressions that can refer to the same object. Moreover, the same utterance – e.g., *the banana* – can be used to refer to different kinds of objects (yellow bananas, brown bananas, etc.). This flexibility poses a challenge for listeners, who have been shown to rapidly draw pragmatic inferences about speakers' referential intentions in online processing. Consequently, understanding how listeners process referring expressions – in particular, to what extent contextual information enters into this process – has been a central topic of psycholinguistic research.

Language is processed incrementally. For instance, eye-tracking experiments have shown that upon hearing an incomplete utterance like *the yellow* in a display like Figure 1a, listeners start to fixate the yellow objects more than other objects in the display even before they hear the disambiguating noun *banana* (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995). [jd: not sure this is the right paper]. But listeners often go beyond the information contained in the signal itself in processing language; they also take into account

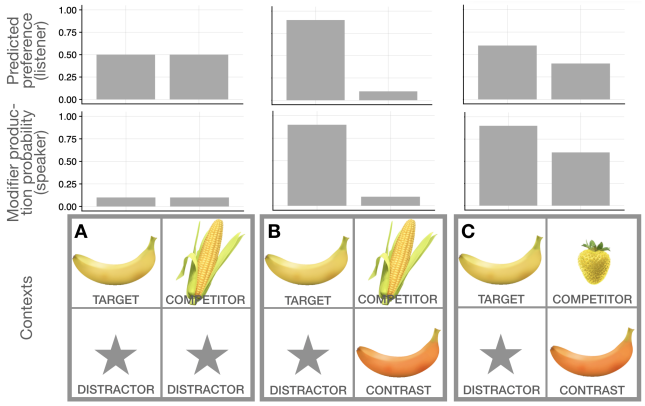

Figure 1: Three contexts, each with a yellow banana as the intended target and another yellow object as its color competitor. The competitor can be typical (in A and B) or atypical (in C), and a contrast can be absent (in A) or present (in B and C). The gray stars represent other distractors that crucially did not share color or shape with any of the other objects in the display.

contextual information – including the nature of other possible referents – to draw rapid pragmatic inferences about a speaker's intended referent [ek: cite]. One such inference that has received much attention in recent years is the so-called *contrastive inference* (Aparicio, Xiang, & Kennedy, 2016; Aparicio, Kennedy, & Xiang, 2018; ?, ?; Grodner & Sedivy, 2011; ?, ?; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). Consider the context in Figure 1b that shows a yellow and an orange banana, a yellow corncob and some other distractor item. When a listener is asked to *Click on the yellow...*, there are two eligible objects to choose from: the yellow banana and the yellow corncob. Rather than consider both yellow objects equally likely target referents, listeners often exhibit a preference, evidenced in increased looks, for the yellow object that has a contrast member of the same type and different color in the display (i.e., the banana, Sedivy et al., 1999; Sedivy, 2003). When the contrast is absent, as in Figure 1a, listeners have no such preference. This preference for the target over the competitor that is elicited by the presence of a contrast (i.e., the orange banana) is considered the result of drawing a contrastive inference.

Contrastive inferences arise as the result of listeners expecting a cooperative speaker to not be more informative than required by the context (?, ?). The presence of a contrast object makes it contextually necessary to include the adjective for establishing reference to the target. In contrast, the adjective is not necessary to refer to the competitor object. Therefore, so the reasoning, listeners reverse-engineer from an observation of the adjective that the intended referent must be the color-congruent object with a contrast member, i.e., the yellow banana in Figure 1b (Aparicio et al., 2016, 2018; ?, ?; Grodner & Sedivy, 2011; ?, ?; Sedivy et al., 1999).

This simple Gricean account that only takes into consideration the contrastive function of the adjective predicts that contrastive inferences should arise whenever the target object occurs in the presence of a contrast object. It is surprising, then, that contrastive inferences are not consistently observed across experiments. While the contrastive inference effect has been replicated reliably in the size adjective domain (Aparicio et al., 2016, 2018; Grodner & Sedivy, 2011; ?, ?; Sedivy et al., 1999), the effect is less stable with color adjectives (Sedivy, 2003). Sedivy (2003) reports that the contrastive inference arises in contexts where the target object has a predictable color (such as the yellow banana in Figure 1) but not when it is replaced by an object with an unpredictable color like a cup, which comes in many colors. They suggest that these objects differ in how likely a speaker is to produce the color modifier for the object in isolation: in the absence of a contrast, a yellow banana is usually just called *the banana* while a yellow cup is often called *the yellow cup*, which Sedivy (2003) calls these objects' *default descriptions*. She suggests that the contrastive inference only arises with objects whose default description does not include the adjective. Only in these cases, she argues, is the observation of the adjective surprising and can be interpreted by the listener as a contrastive signal.

[jd: this paragraph can read: "In addition to expectations of informativity as described above, contrastive inferences have been proposed to depend on the semantics of the adjective involved, such that only relative adjectives (e.g., size adjectives) but not absolute adjectives (e.g., color adjectives) result in the inference (Aparicio et al., 2018). Furthermore the effect only arises when the listener considers the speaker as a reliable speaker who does not violate communicative norms (Grodner & Sedivy, 2011; ?, ?)."][ek: Yes, that sounds good, but doesn't Aparicio:2018 claim that color adjectives show the contrastive inference? The paper says: "Color adjectives, relative adjectives and maximum standard absolute adjectives were rated as overinformative when used as modifiers in the absence of contrast, and gave rise to RECs; minimum standard absolute adjectives were not rated as overinformative when used as modifiers in the absence of contrast, and did not give rise to RECs" (RECs = referential effects of contrast)][jd: true, they're the only ones who have ever found it the other way around. you can state it more neutrally in terms of the proposed semantics mattering, but don't state the direction (or say that different people have made the claim in different directions.) crucially, don't get bogged down in the details.]

In this paper, we provide a novel account of contrastive inference that has the potential to unify the above properties by reducing them to listeners' expectations about the speaker's contextual probability of producing the pre-nominal adjective. In so doing, we follow recent research highlighting the importance of listeners' generative model of the speaker in generating pragmatic inferences (?, ?, ?, ?, ?)[jd: you can really cite a bunch of rsa papers, and i'm tracking down the relevant non-pragmatics references, too]. We propose to formalize the relevant listener-side reasoning within the Rational-Speech Act (RSA) framework (Frank & Goodman, 2012; Goodman & Frank, 2016), a state-of-the-art computational framework that models pragmatic inference as the result of listeners performing Bayesian inference on their speaker model and their prior beliefs about likely meanings, thereby giving the speaker model a central role in the inference. It provides a way to quantitatively assess the probability that a listener with prior beliefs and expectations about the speaker assigns to possible referents after observing partial sentences of the form *Click on the yellow...*. This account shifts the explanatory focus away from specific cognitive and linguistic factors that influence contrastive inference and towards listener's production expectations (and their prior beliefs, which we don't treat in depth in this paper).

For this investigation it is important to distinguish between two notions: the theoretical construct of a *contrastive inference* and the behavioral pattern that manifests as a *target preference*, i.e., a preference for the target over the competitor. A target preference in contexts in which a contrast is present has been interpreted as evidence for a contrastive inference [jd: cite?][ek: I'm not sure what to cite here. As far as I understand it, all papers do the statistical analysis on the basis of the contrast-present and contrast-absent condition, even though they then sometimes just talk about *target preference*]. Though a contrastive inference may result in a target preference, it need not. It may also manifest as an increase in target probability (but no resulting target *preference*) in a context that includes a contrast compared to a context where the contrast is absent. That means there can be a contrastive inference even in the absence of a target preference, e.g., if there was a target dispreference in the absence of the contrast. Conversely, a target preference may be observed for reasons other than the presence of a contrast – we discuss this in more detail below. [jd: i wonder if the middle of this paragraph is necessary in this much detail – will it help with understanding anything later on? i thought we just need this paragrahp to introduce the notion of a target preference independently of contrastive inference ,because it's the target preference across contexts that we'll mostly be focusing on in the rsa analysis]

In this paper we will first show that our production-centric account makes the same qualitative predictions about the basic contrastive inference effect as for instance the default description account (Sedivy, 2003). We then derive new pre-

dictions about the size of target preferences across different contrast-present and contrast-absent contexts. We report a free production study we conducted to elicit modifier probability estimates, which will be used to determine quantitative model predictions. For the evaluation of those predictions, we compare them to empirical comprehension data which we elicited using an incremental decision task.

A Bayesian account of contrastive inference

[ek: "Likelihood" and "probability" are all scrambled up here. Don't forget to fix that!] [jd: that didn't strike me anywhere, and i was looking out for it]

[jd: this first paragraph seems a little out of the blue. i understand you want to introduce the default description account – perhaps you can just say that previous accounts of contrastive inference have ascribed more or less relevance to the speaker, and that the arguably most comprehensive account of at least the difference between the predictable/unpredictable-color result is the default description account?] In the literature, different factors have been considered to give rise to contrastive inference which indirectly put more or less relevance onto the speaker. In the *default description* account for instance (Sedivy et al., 1999), a speaker is only considered as to the creation of the default descriptions and is completely independent of the context the target is presented in.

[jd: the following already assumes too much on the part of the reader. most people still don't know what rsa is, say a little more in prose about what rsa does – can take inspiration from degen et al in press or goodman and stuhlmueller 2013]

The Rational Speech Act framework gives the speaker a central role in pragmatic interpretation since the listener P_{L_1} determines the probability of each possible referent r being the target via Bayesian inference about two quantities: the speaker's possible utterances u for each r in context C $P_{S_1}(u|r,C)$ and the listener's prior beliefs about possible referents $P(r|C)$.

$$P_{L_1}(r|u,C) = \frac{P_{S_1}(u|r,C) * P(r|C)}{\Sigma P_{S_1}(u|r_i,C) * P(r_i|C)} \quad (1)$$

[jd: something is wrong with the math, what's being summed over in the denominator? i would instead just show equation 2 but add the prior back in and just say in prose that we won't consider it because we assume it to be uniform (and add footnote saying that the empirical results justify that assumption, point reader to exp 2)]

To simplify the following example, we will assume that listeners have a uniform prior $P(r|C)$ over all objects in the display. Then the RSA model predicts a direct relationship between the production probabilities P_{S_1} and the listener's distribution over possible referents P_{L_1}.

$$P_{L_1}(r|u,C) \propto P_{S_1}(u|r,C) \quad (2)$$

While RSA has typically been applied to the analysis of full utterances, it can straightforwardly be extended to generate predictions at the sub-sentential level. To receive RSA predictions for an incomplete referring expression such as *the yellow...*, we take P_{S_1} to correspond to the contextual probability of color mention for each referent in the display. Let's investigate the qualitative predictions this account makes:

Consider the example contexts in Figure 1. Upon hearing the modifier *yellow*, the pragmatic listener P_{L_1} considers how likely a speaker is to include this modifier in their referring expression for each object in the display. Since only the target (yellow banana) and the competitor (corncob) are yellow, we assume that the production probabilities of *yellow* for the other objects in the display are 0. This only leaves the target and the competitor as potential referents.

Hypothetical modifier production probabilities for target and competitor are displayed in the middle row of Figure 1.Assume that in the absence of a contrast object (Figure 1a), speakers are equally unlikely to include the color modifier when referring to the target banana (probability 0.1) and its color competitor, the corncob (0.1). To obtain pragmatic listener predictions, these probabilities are simply renormalized, resulting in a target preference of 0.5, i.e., the pragmatic listener does not prefer one potential referent over the other.

Does RSA predict the contrastive inference effect in context Figure 1b? Assuming that the presence of the contrasting orange banana does not affect the speaker's modifier production probability for the competitor corncob but does increase modifier production probability for the target banana to 0.9, renormalizing the production probabilities results in a target preference of 0.9 – thus reproducing the classic contrastive inference.

In contrast to previous accounts of contrastive inference, modifier production probabilities are expected to directly drive the contrastive inference and associated target preference. Since the contrastive inference is the difference in target preference between contrast conditions and the target preference depends on the modifier production probabilities of the target and the competitor, the competitor takes on a central role in these predictions.[jd: this seems to contradict what is said in the intro paragraph about the relation between target pref and contrastive inference] This suggests that increasing the modifier production probabilities for the competitor should lead to a decrease in target preference. It has been established that speakers are more likely to include color modifiers in referring expressions for objects in isolation when they appear in an atypical than in a typical color (?, ?, ?, ?). It is thus likely that the atypical yellow strawberry in Figure 1c would be more likely to elicit a color mention than the typical corncob in Figure 1b. Assuming a modifier production probability of 0.6, this contrast-present context shows a much smaller increase in target preference compared to the contrast-absent context. In other words, the size of the contrastive inference is predicted to be smaller with an atypical compared to a typical competitor, keeping target typicality and contrast presence constant. This predicts that the size of

the contrastive inference can vary depending not only on features of the target (as previously shown by Sedivy, 2003)[jd: and rubio-fernandez?], but also crucially on features of the color competitor.

[jd: these last few paragraphs that bring in typicality and go through predictions are very nicely done]

To investigate this novel prediction, we first elicited modifier production probabilities (i.e., an estimate of $P_{S_1}(u|r,C)$) in a free production interactive reference game (Exp. 1) in contexts that varied in the presence of a contrast, the typicality of the target, and the typicality of the competitor. This allowed us to generate pragmatic listener probabilities for each display. We then evaluated model performance by comparing these predictions to empirically elicited interpretations (Exp. 2).

Experiment 1: Modifier Production in an Interactive Reference Game

Experiment 1 was aimed at obtaining modifier production probabilities for all the displays ultimately used in the contrastive inference experiment (Experiment 2). We elicited these probabilities in a free production interactive reference game. We expect that the typicality of a color for an object will affect these modifier production probabilities, i.e., we expect speakers to call a yellow banana simply *the banana*, but an orange banana *the orange banana*. We take the results as the modifier production probabilities a listener can expect for each object.

Participants

We recruited 282 participants over Amazon's Mechanical Turk, who were randomly matched to form listener-speaker chat pairs (i.e., 141 pairs in total). Each participant was paid \$2.30 (approximately \$11-\$14/hr)[1]. We restricted participation to workers with IP addresses in the US and an approval rate of previous work above 97%. We excluded 29 participant pairs because of multiple participation and for primarily using unnatural descriptions such as *should be yellow, must have teeth to eat* for *red corn*.

Material

Each context included four items, as displayed in Figure 2. The pool of items consisted of 10 types (banana, broccoli, carrot, corn, egg, lettuce, pumpkin, strawberry, swan, tomato), each of which could occur in a typical and atypical color. For example, the broccoli could occur in its typical color green or in the atypical color red. The resulting pool contained 20 items, 10 of which were atypically colored. The number of colors were counterbalanced such as each color occurred twice as a typical and twice as an atypical instance.

[1]The experiment was preregistered on [ek: id]. Originally, we recruited 68 participants and then ran a follow-up with 214 more to get enough data for the evaluation of the RSA model. The results from the first 68 participants do not differ from the full data set, which is why we present them collapsed.

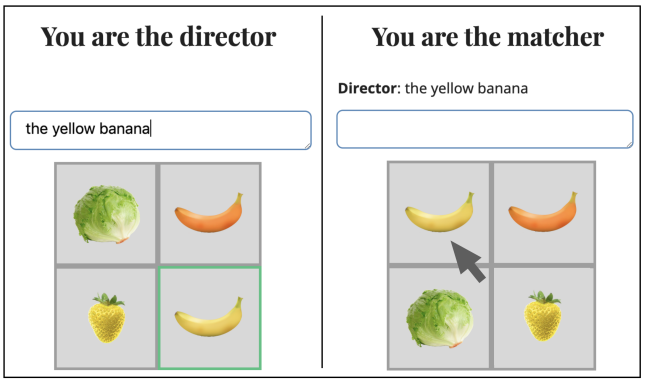

Figure 2: Example display for the interactive reference game (Experiment 1). Both, the speaker (here *Director*) and listener (*Matcher*) see the same four objects but in a scrambled order. Additionally, the speaker sees a green border around one of the objects, marking the intended target which the listener needs to select.

All items were carefully normed for color-diagnosticity (?, ?), typicality and nameability.

Design

The contexts varied in the typicality of the target, the typicality of the competitor and the presence of a contrast, resulting in eight conditions. We needed to elicit the modifier production probabilities for the target and the competitor. In contexts where the contrast was absent, this distinction is irrelevant. For example, when target and competitor are both (a)typical, either could be underlyingly coded as the target. Similarly, the modifier production probability for a typical target with an atypical competitor is the same as the probability for a typical competitor in a context with an atypical target. In contrast-present contexts, the target-competitor distinction matters, which is why speakers had to communicate the competitor half of the trials and the target in the other half.

The fillers were eight randomly created contexts where the contrast or the distractor had to be communicated. Overall, each participant saw 60 different contexts (32 critical trials) in a completely randomized order.

Procedure

Participants were randomly paired up and each was randomly assigned either to the role of a speaker or listener. They could communicate freely through a real-time multi-player interface similar to (?, ?). The speaker was instructed to communicate a target object out of a four-object context to the listener. The target could be identified by a green border surrounding it. The speaker and the listener saw the same set of objects but in a randomized order to avoid trivial position-based references such as "the left one". After the listener clicked on the presumed target, both the speaker and listener received feedback about whether the right object had been selected.

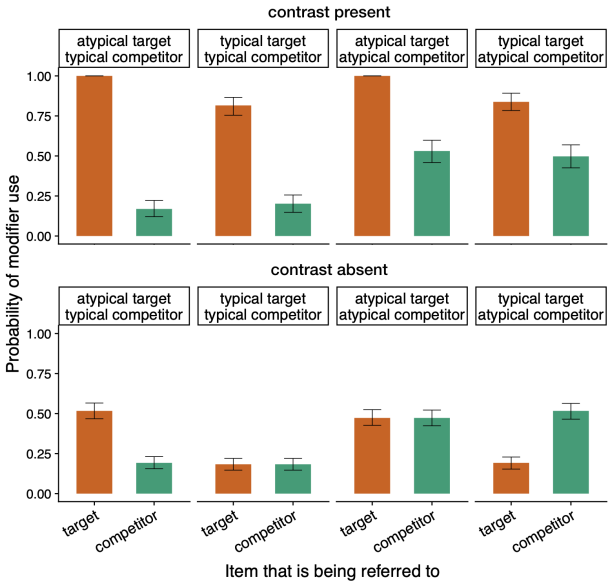

Figure 3: Results of the production study, showing the probability of modifier use for the target and competitor in each condition. Error bars are 95% bootstrapped confidence intervals.

Results

Figure 4 shows the probability of color modifier mention for the target and competitor in each condition[2].

When a contrast to the target is present (e.g., another banana), a speaker needs to include the color modifier to fully disambiguate the two items (see the upper row in Figure 4). When the target was atypical, speakers always used the color modifier, while this was not always the case for a typical target.

When the contrast was absent (see the lower row in Figure 4), speakers were more likely to include a color modifier when referring to an atypical target than a typical one.

Independent of contrast, speakers were more likely to include the color modifier for an atypical color competitor over a typical one.

The results of this production experiment show that the probability of a speaker's modifier use is modulated by the color typicality of the item and the presence of a contrast. Our experiment therefore successfully manipulates the modifier production probabilities a listener can expect in different contexts.

Comprehension Experiment: An Incremental Decision Task

To investigate which objects listeners consider to be the most likely referent after observing the color adjective, we conducted an incremental decision task (?, ?). This is an offline task to investigate a listener's belief about the intended referent while gradually unfolding the referring expression.

Participants

We recruited 239 participants over Amazon's Mechanical Turk, 121 of which saw atypical color competitors and 118 saw typical color competitors in the critical trials[3]. Each of them were paid $1.80 for their participation (10-16/hr). We restricted participation to workers with IP addresses in the US and an approval rate of previous work above 97%. 27 participants were excluded because they indicated that they did the experiment incorrectly, English was not their native language, or they gave more than 20% erroneous responses[4]. 211 participants remain, 108 of which were in the atypical competitor and 103 were in the typical competitor condition.

Material

The item pool is the same as in the production study (Experiment 1).

Design

Participants completed 55 trials in total, 20 of which were critical trials and 35 were fillers. The contexts varied for each participant with respect to the presence of a contrast and the target's color typicality (within-subject manipulation). Participants were randomly assigned to see either typical or atypical competitors on critical trials (between-subject manipulation). All critical trials included color modified utterances. To avoid learning effects, we included filler trials with unmodified referring expressions and with referents other than the targets of the critical trials.

Procedure

This experiment is a one-player adaptation of the production study explained above and follows the design of an incremental decision task (?, ?).

All participant were assigned the role of the listener, which means that they needed to identify which object was the target given a referring expression placed above the context. Crucially, the referring expression was only gradually revealed and participants had to choose an object each time before the trial continued. In each critical trial, three choices had to be made: (1) before receiving any information about the referent (i.e, after observing "Click on the"), (2) after receiving the adjective ("Click on the yellow") and (3) after receiving the full referring expression with the disambiguating noun ("Click on the yellow banana!").

To center the position of the mouse after each selection, a button appeared in the center of the grid which had to be clicked to reveal the next word or to advance to the next trial.

[2]Note that some data is duplicated in the conditions where the contrast is absent (as described in *Design*)

[3]The experiment was preregistered on [ek: id]. Originally, we recruited 80 participants and then ran a follow-up with 140 more to get enough data for the evaluation of the RSA model. The results from the first 80 participants do not differ from the full data set, which is why we present them collapsed.

[4]An erroneous response is defined as a selection of a non-target object after observing the fully disambiguating noun.

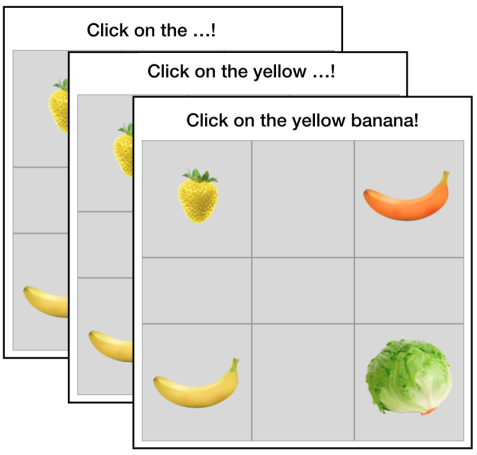

Figure 4: Design of the incremental decision task. The referring expression was placed above the grid and revealed gradually. After each new word participants made a selection indicating their best guess about the intended target.

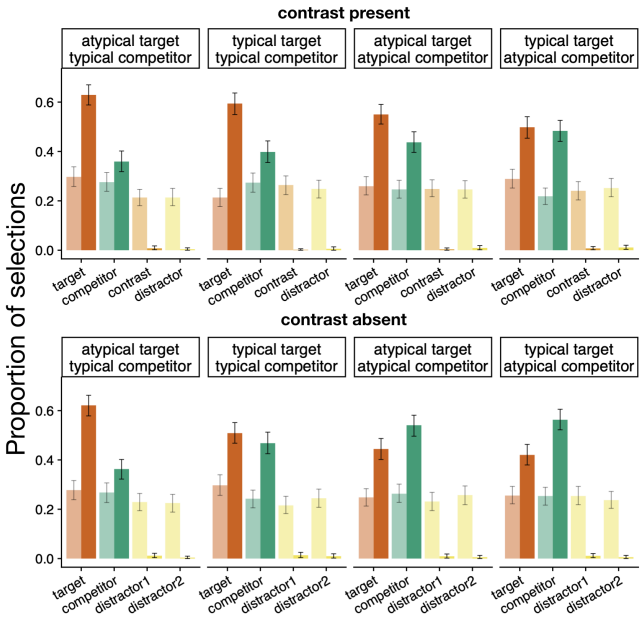

Figure 5: Results for the comprehension study, showing the proportion of selections for each item in the display and each condition. The bars in lighter colors indicate the selections before, the darker bars are the selections after the adjective was observed. Error bars are 95% bootstrapped confidence intervals.

Trials were randomized with the only restriction that modified utterances that referred to a typical object with no contrast only appeared after the 15th trial to minimize the risk that the speaker was perceived as unreliable (Grodner & Sedivy, 2011).

Before participants proceeded to the main trials, they had to complete four practice trials constructed from the speaker perspective, which were introduced to familiarize the participants with the task.

Results

Figure 5 shows the proportion of object selections before reading the adjective (lighter colors) and after the adjective (darker colors), grouped by context condition. Before an adjective is observed, all items should appear equally likely to be the target, which is supported by the generally uniform distribution in all conditions. After the adjective is revealed (darker colors), only the target and competitor are legible options and we predicted that the presence of the contrast and the typicality of the objects will affect the listeners' object choices.

When the contrast is present (upper row in Figure 5), there is a general preference for target over competitor selections. This preference is biggest for the case when the target is atypical and the competitor is typical and disappears for when the target is typical and the competitor is atypical.

When the contrast is absent (lower row in Figure 5) and target and competitor differ in typicality, there is a preference for the item with the atypical color. When the two items share their typicality, the selection are approximately at the same rate.

Qualitative post-hoc analyses revealed that the selections after reading the adjective were also affected by the participant's previous selection. A participant who previously selected the competitor was more likely to select the competitor again than switch to the target (and vice versa). But since the object selections before the adjective occurred are uniformly distributed, any patterns that appear after the adjective cannot be an artifact of the reselection bias.

These results clearly show that the color typicality of the objects in the display affect the inference listeners draw about the intended referent. An atypical competitor alone can promote the competitor over the target when the contrast is absent and can even make the target preference disappear when a contrast is present. It is therefore highly relevant to control for the quality of the competitor when assessing contrastive inferences. Choosing an atypical target makes the contrast-present and contrast-absent conditions more similar, suggesting a smaller contrastive inference. This replicates the finding that the contrastive inference did not appear with items of unpredictable colors (Sedivy, 2003).

Model evaluation

To assess the relationship between the modifier production probabilities and the comprehension data in the simplest way, we will assume a flat prior over all objects in the display. This choice is further justified by the uniform distribution over all objects before receiving any information about the target in the comprehension experiment. The probabilities to choose the target over the competitor are then the normalized modifier production probabilities as shown in Equation (3), where r is the possible referent, u the utterance and C the specific

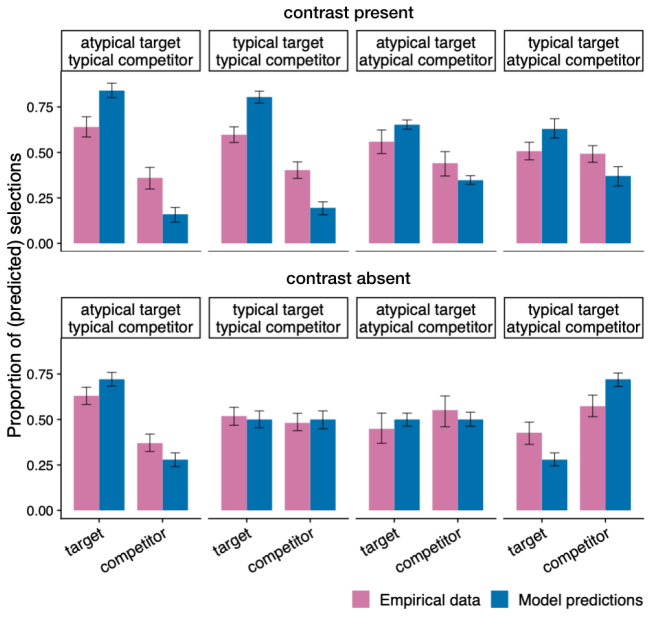

Figure 6: This is a figure.

context.

$$P_{L_1}(r|u,C) = \frac{P_{S_1}(u|r,C)}{P_{S_1}(u|r_{target},C) + P_{S_1}(u|r_{comp},C)} \quad (3)$$

Figure 6 shows the model predictions (in blue) and the empirical results (in purple) for target and competitor selection after observing the adjective.

Using the modifier production probabilities obtained in Experiment 1, the model qualitatively predicts the patterns for the different context conditions.

Quantitatively the RSA model is a significant predictor for the empirically elicited comprehension data ($E = 1.46, CI = [1.01, 1.91]$)[5] and its predictions highly correlate with the empirical results ($r = 0.91$). However, it generally predicts more extreme probabilities than are borne out in the empirical data, as shown in Figure 7. The model overpredicts target selections in high target preference conditions and underpredicts target selections in low target preference conditions.

One possible explanation is that the inferences appear smaller in the comprehension study because of the bias to reselect the previously selected object (as described in the results of Exp. 2). If a participant observes an adjective that could elicit a contrastive inference but the participant selected the competitor in the prior window, the non-switching bias counteracts the contrastive inference. This can explain why the range of empirical target selection proportions is compressed towards chance. Whether this is indeed an artifact of the incremental decision task could be investigated by rerunning the experiment without a prior window selection.

[5]Results of a Bayesian mixed effects regression model $target\_selection \sim RSA\_prediction + previous\_selection + (1 + RSA\_prediction|participant)$.

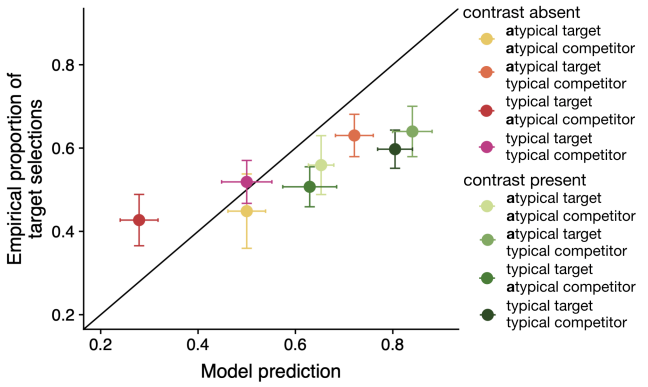

Figure 7: This is a figure.[ek: put contrast-present up]

Similarly, in an eye-tracking version of the experiment, where participants do not make explicit looking decisions, the non-switching bias is likely to be weaker.[jd: i'm actually not entirely sure this is true, given that the best predictor of fixation at time point t0 is fixation at time point t-1, but we can let it stand]

Overall, these results suggest a strong connection between referring expression interpretation and production. Only using the probability of encountering the observed adjective, the RSA model can qualitatively and quantitatively predict the empirically elicited comprehension data. Although we replicate that the contrastive inference can be elicited in an offline incremental decision task, the model results suggest that the selection biases in the paradigm might reduce the size of the inferences. We expect this bias to reduce in an eye-tracking paradigm, which is an immediate future direction for this work.

Discussion

In this paper, we provided a novel account of contrastive inference in which we argue for a speaker-centric model of comprehension. We use the Rational-Speech Act (RSA) framework to make quantitative predictions about the behavior a pragmatic listener *should* exhibit when provided with different contexts. This account shifts the focus away from specific cognitive and linguistic factors that have been discussed to affect contrastive inference in the literature onto listener's production expectations (and their prior beliefs). We show that this speaker-centric model cannot only account for the general case of contrastive inference, but makes further predictions onto why contrastive inference appears to be less stable with color adjectives.

In contrast to previous accounts, it is not simply the production probability of the target that matters (as suggested in a default description account (Sedivy, 2003)), but instead the relative modifier production probability of *all* objects in the display, which are then evaluated against each other. In this particular case this means that it assigns a central role to the color competitor in the display. The empirical results confirm

that the choice of the competitor affects the interpretation of the utterance, providing evidence for this highly pragmatic account of comprehension.

We assessed the model predictions by collecting object selections in an incremental decision task (?, ?), replicating that it generally can elicit the contrastive inference. Crucially the results show high variation between context conditions dependent on the typicality of the target and competitor. In other words, by varying the modifier production probabilities, we can make the contrastive inference appear strong and almost make it disappear. This range provides a challenge for accounts that argue for a uniform quality about adjective semantics that affect contrastive inference. A speaker-centric account predicts instead that the general differences observed between different types of adjectives is in fact mediated by how likely a listener expects the modifier to be produced and an interesting new avenue for future investigations.

References

Aparicio, H., Kennedy, C., & Xiang, M. (2018). Perceived informativity and referential effects of contrast in adjectivally modified nps. In *The semantics of gradability, vagueness, and scale structure* (pp. 199–220). Springer.

Aparicio, H., Xiang, M., & Kennedy, C. (2016). Processing gradable adjectives in context: A visual world study. In *Semantics and linguistic theory* (Vol. 25, pp. 413–432).

Degen, J., Hawkins, R. D., Graf, C., Kreiss, E., & Goodman, N. D. (2019). When redundancy is useful: A bayesian approach to 'overinformative' referring expressions. *Psychological Review*.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of psycholinguistic research*, *24*(6), 409–436.

Frank, M. C., & Goodman, N. D. (2012, May). Predicting pragmatic reasoning in language games. *SCIENCE*, *336*, 1.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.

Grodner, D., & Sedivy, J. C. (2011). 10 the effect of speaker-specific information on pragmatic inferences. In *The processing and acquisition of reference* (Vol. 2327, pp. 239–272). MIT Press.

Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of psycholinguistic research*, *32*(1), 3–23.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*(2), 109–147.