# Availability and Timing of Informativity Inferences

**Alice Rees (al.b.rees.13@gmail.com)**
**Hannah Rohde (hannah.rohde@ed.ac.uk)**
School of Philosophy, Psychology, & Language Sciences
University of Edinburgh, EH8 9AB

**Abstract**

Conversational partners expect each other to communicate rationally and cooperatively and to contribute relevant and informative utterances. Occasionally, however, speakers produce trivial utterances which may violate our expectations of informativity. These utterances can prompt listeners to draw inferences about a speaker's goals in producing such an utterance. Here we present two studies investigating how and when listeners derive *informativity-based* inferences. The results demonstrate that speaker knowledge plays an important role in the computation of inferences. Furthermore, the timecourse of the results suggests that these inferences do not always arise automatically and that their computation is costly.

**Keywords:** pragmatics; speaker knowledge, inferences

## Introduction

Language is a tool that we use to achieve many goals. For example, to inform, persuade, or surprise. Part of what listeners do in communicative interactions is to determine the purposes or intentions underlying what speakers say (Grice, 1975; Levinson, 2000). On the surface, communication seems simple enough: as speakers we decide what it is we want to convey to our interlocutor and then we produce a signal to achieve that; as listeners we take the incoming signal and map the words onto meanings. Indeed sometimes language is used directly (or transparently), such that a speaker's utterance and the intended meaning go hand in hand. In some cases, however, language is used indirectly (non-transparently) whereby an utterance does not transparently map onto the intended meaning, and successful communication requires listeners to go beyond the transparent and infer additional meaning. This potential indirectness presents listeners with a challenge, namely when to draw an inference? The present work examines so-called *triviality* or *informativity-based* inferences—inferences that arise when the asserted meaning of a speaker's contribution fails to satisfy listener expectations for cooperativity. Specifically, contributions falling short of providing sufficiently newsworthy content invite the listener to speculate about the speaker's communicative goals and the possible intended meaning, which may go beyond what was explicitly said.

As competent language users we are expected to behave cooperatively as rational communicators who use language to contribute relevant and informative content to a discussion (e.g. Bohn, Tessler, & Goodman, 2019; Brown & Dell, 1987; Grice, 1975; Rohde, Futrell, & Lucas, 2021). These expectations are not always upheld, however. Speakers often violate the expectations of their partner, for example providing underinformative or overinformative (and potentially redundant) utterances. Consider hearing your conversational partner say "I ate some of the cookies" or that "the soup is warm". Both of these utterances license the listener to make inferences. In the first instance, that the speaker ate *some but not all* of the cookies and in the second that the soup is *not hot*. These inferences are examples of scalar implicatures which arise when a speaker chooses not to use a stronger expression. Since *some* and *hot* are members of lexical scales ordered on informativity (<some, all> and <warm, hot>), when a speaker uses a less informative expression from the scale, listeners try to reconcile their expectations regarding speaker informativity by inferring that the stronger (more informative) expression doesn't hold, thereby deriving a scalar implicature (Gazdar, 1979; Geurts, 2010; Grice, 1975; Horn, 1972; Levinson, 2000). An important characteristic of inferences such as these is that they are optional; that is, listeners are never obliged to draw an inference, which begs the question of whether and when listeners do draw such inferences.

It is not the case that inferences arise only if a speaker is underinformative; they can also arise if a speaker has been overinformative in their utterances. Seemingly redundant or overly informative utterances are common (Baker, Gill, & Cassell, 2008; Engelhardt, Bailey, & Ferreira, 2006; Gann & Barr, 2012; Grice, 1979; Kravtchenko & Demberg, 2022; Levinson, 2000; Wardlow-Lane, Groisman, & Ferreira, 2006). For example, requesting a "yellow banana" may be considered overinformative or redundant since prototypical bananas are yellow, i.e. *yellow* is a common feature of *banana,* and therefore it is usually unnecessary to specify *yellow*. Consequently, listeners may search for a reason, or communicative goal that would justify the inclusion of *yellow*. For example, listeners may infer there is another non-yellow banana in the context that would make uttering "banana" alone insufficient (Sedivy, 2003; c.f. Rubio-Fernandez, 2016; 2019). Another possibility is that, for the speaker, *yellow* bananas are of particular interest and therefore the colour is worth mentioning (Westerbeek, Koolen, & Maes, 2015) or that the mention of the colour is particularly helpful to the listener for locating the object in a visual scene (Long, et al., 2021; Rubio-Fernandez 2016; Wu & Gibson, 2021). Indeed, listeners often infer that a speaker's mention of an object feature reflects that features's atypicality (Bergey & Yurovsky, 2022; Horowitz & Franks, 2016; Kreiss & Degen, 2020).



In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.* 

In the examples discussed above, a listener need not compute the inferences outlined. The utterances could be taken as descriptive, with low informativity. The listener must establish if the speaker should be interpreted as using direct (transparent) language, which makes no additional demands on the listener, or indirect (non-transparent) language, which would require listeners to compute further meaning. Although a speaker is never on record as having asserted the additional content, a listener may nonetheless be inclined to identify and infer such meaning because doing so may be necessary for maintaining the sense and relevance of a conversational contribution (see also, Relevance Theory, Sperber & Wilson, 1995; Carston, 2004).

Typically, inferencing is studied with particular emphasis on narrow classes of words that generate scalar implicatures (van Tiel et al., 2016). However, many of the inferences that arise during the course of interaction are not limited to scalar terms. In the present work we branch out from traditional investigations of scalar implicatures to consider a broader class of inferences that can arise during a conversation. Particularly, we focus on the types of inference that arise when utterances fail to satisfy our expectations for conversationally appropriate levels of informativity.

## Informativity inferences

The computation of an inference can be prompted both by a speaker's choices regarding what content to mention and how to package that content. Consider Levinson's (2000) M-principle, *what is said in an abnormal way is not normal,* and I- principle, *what is not said is the obvious.* Under the M-principle "Bill caused the car to stop" implies that the car was stopped in a non-stereotypical way (i.e. not using the brakes). Choosing to describe the event in this way is unconventional, thereby rendering the utterance marked, which can lead listeners to try to reason why the utterance was produced in an unconventional manner. For example, they may infer that the event itself was atypical. Under the I-principle, we do not expect speakers to state the obvious or tell us information that is easily inferable. Consequently, saying that "Bill stopped the car" may imply that the car was stopped stereotypically i.e. Bill used the brakes. However, it may further imply that there was a particular reason for Bill to stop the car (it is not usually noteworthy to mention that a car was stopped unless the object of interest is the cause i.e. Bill stopped the car because the police flagged him down). That is, the choice to produce the utterance at all, given the obvious nature of the content, can be signal enough for a listener to compute an inference. These kinds of particularised inferences go beyond the conventionalised classes of implicatures traditionally studied and are wide ranging.

Kravtchenko & Demberg (2022) demonstrated that when speakers produced redundant event descriptions (e.g. mentioning paying for items in a shop), participants interpreted these utterances as indicating atypicality of the event (i.e. participants inferred that the character does not usually pay for items). This can be explained through the I-principle. The script knowledge of going shopping includes paying, thus it is redundant to include this information (I-principle). In order to reconcile the violation of conversational norms, listeners reasoned that this behaviour was atypical, based on their real-world knowledge. In the present work we consider whether this reasoning extends to other types of utterances where the listener doesn't rely on their own knowledge of the world but instead on their understanding of the knowledge state of the speaker.

Imagine a listener who encounters (1) below out of the blue – a simple assertion that the walls are blue.

(1) "The library walls are blue"
  a. → *the situation has changed*
  b. → *the walls used to be different*

The triviality of (1) may invite listeners to reason about why the speaker chose to utter this, what the speaker's goals were in doing so, and potentially what additional meaning can be inferred since a goal of being informative seems implausible for the asserted content alone.

Additional meaning can arise if listeners believe the speaker is: (i) knowledgeable of the situation over time and (ii) sufficiently cooperative to adhere to the general goal of conveying sufficient information for the current exchange. If these assumptions hold, one way of reconciling the production of an utterance about a fairly typical or trivial situation (henceforth "trivial utterance") is to infer (1a), whereby what is informative is not the colour of the walls but the (less typical) event of their having been painted a new colour. In other words, (1) permits the possible inference of (1b). We refer to this kind of inferred meaning as a type of *informativity-based inference*, one that arises when an utterance's direct meaning fails to satisfy listener expectations for cooperative levels of information exchange. The goal of the present study is to test how speaker knowledgeability affects the computation of *informativity-based* inferences and how readily these inferences are computed.

A central question regarding pragmatic inferences relates to whether or not the process of inferencing is costly. A number of studies argue that scalar implicatures are costly to generate (e.g. Breheny Katsos Williams 2006; Bott & Noveck, 2004; Huang & Snedeker, 2009, 2018; Noveck & Posada, 2003), in keeping with a literal-first perspective of implicature computation whereby listeners first compute the literal interpretation and then go on to derive the implicature. If *informativity inferences* are processed in a similar manner to scalar implicatures, a literal-first perspective would predict a cost for the computation of such inferences as reflected in processing times. However, scalar implicatures are not always costly to compute and have been speculated to be computed as easily as a literal interpretation (e.g. Breheny Ferguson Katsos 2013; Grodner, Klein, Carbary & Tanenhaus, 2010). It has been suggested that the processing cost observed is not the same across all implicatures, and that inferences are drawn with differing levels of ease, dependent on a number of constraints (e.g. Speaker knowledge; Bergen & Grodner, 2013; Kampa & Papafragou, 2020; Papafragou, Friedberg, & Cohen, 2018; context, Breheny, Katsos, & Williams, 2005; Degen & Tanenhaus, 2015; Singh, 2021; van Tiel, Pankratz, & Sun, 2019).

In the present work we focus on how speaker knowledge affects the processing of *informativity-based* inferences (c.f. Bergen & Grodner, 2013; Kampa & Papafragou, 2020; Papafragou, Friedberg, & Cohen, 2018). We demonstrate that *informativity* inferences are computed more often when trivial utterances are produced by a knowledgeable speaker. Based on reaction time data there appears to be a processing cost associated with computing such inference. However, this claim is tentative as data collection is ongoing.

## Experiment overview

To assess the computation of *informativity-based* inferences, we manipulated speaker knowledgeability by varying whether the speaker described a location they were familiar with or a location they were unfamiliar with. When confronted with utterances such as "the library walls are blue", if listeners reason about why a speaker may produce such a trivial utterance, one possible reconciliation is the inference that this is not the usual state of affairs (via the I-principle). As outlined above, this reasoning only follows if the speaker is knowledgeable about a location over time. To assess if participants compute *informativity-based* inferences we ask participants about the situation (e.g. the colour of the library walls) at a previous point in time. Thus we predict that there will be greater rates of inference computation when a speaker is knowledgeable than unknowledgeable.

In addition, we ask how readily informativity-based inferences are derived; are they computed as a part of normal comprehension processes, i.e. as soon as a listener encounters a trivial utterance, is the mismatch in expectations and reality reconciled? Or are these inferences too costly to compute unless the listener is specifically prompted to consider the speaker's goals? If *informativity-based* inferences arise immediately upon, or very soon after, encountering a trivial utterance, the inference computation would be predicted to incur an immediate cost, observable in longer reading times when processing the utterance itself. However, these inferences may not arise automatically, in which case there may be a delay at a later point when the inference is specifically prompted.

## Experiment 1

### Method

#### Participants

We recruited N=200 participants, who reported being fluent in English, from the crowdsourcing site Prolific. Two participant failed to complete the experiment and were removed from all analyses.

#### Design & Materials

To manipulate Speaker Knowledgeability, participants were presented with a speaker Suzy, a child who is telling her dad about her day at either a place she is familiar with (school) or somewhere she is unfamiliar with (Prime Minister's office; PM). In total there were 20 critical utterances about plausibly changeable situations, for example "I saw that the library walls are blue". Since walls can be painted, it is plausible that the situation being described may have changed (see osf [1]for complete list of stimuli). We chose situations such as these rather than situations that are highly unlikely to change (e.g. There are white lines at the zebra crossing) or situations that are highly changeable (e.g. There are puddles on the ground).

Speaker Knowledgeability was manipulated within participants so that each participant saw 10 utterances for the location familiar to the speaker and 10 for the unfamiliar location. Presentation of the utterances was counterbalanced across participants such that across all participants all utterances were seen in both conditions but each participant only saw an utterance once in one condition. Participants were asked "was it the same a few months ago" and had to respond true or false. There were an additional 10 filler items which included an attention check.

#### Procedure

The experiment was hosted and administered online through Psytoolkit (Stoet, 2010; 2017). Item presentation was blocked and counterbalanced to ensure an equal number of participants saw the familiar and unfamiliar location first and that trials within each block were fully randomised. At the start of a block participants were given an introductory slide introducing Suzy, her dad, and the location they were going to be talking about. To progress through the experiments participants were required to press a button. Participants were not given any instructions about the speed with which they should respond. On critical trials following presentation of the sentence, on a separate screen which participants moved to through a button press, participants were asked "Was it the same a few months ago?" and could respond either "yes" or "no". Responding "no" indicates (or is at least compatible with) participants' derivation of an *informativity* inference. Reading times were calculated as the time spent on the critical trial where the utterances were presented. Response times were measured as participants time to respond to the question on the subsequent screen.

### Results

All analyses were carried out in in R (Version 4.0.3, R core team, 2020) using lme4 (Version 1.1-23; Bates, Mächler, Bolker, & Walker, 2015) and emmeans (Version 1.6.0; Lenth et al. 2021). We always used the maximal model that allowed for convergence. We analysed the binary responses (yes/no) using a logistic regression with Speaker Knowledge as the

[1] https://osf.io/nbhya/?view_only=1a87c27512bf406c9fcaa4c86e08764c

predictor. We analysed the log transformed response times and reading times for critical sentences using a regression model with Speaker Knowledge and Response type as predictors. Variables were recoded such that, for Speaker Knowledgeability, the unfamiliar PM office was coded as -0.5 and the familiar school as 0.5 and for response type "No" was coded as -0.5 and "Yes" was coded as 0.5.

**Inference rates.** As seen in Figure 1, there was a greater rate of "no" responses in the familiar school condition. This pattern is borne out in the model which shows a main effect of Speaker Knowledge (β=.976, SE=.083, z=-11.743 $p < .001$); participants derived more *informativity* inferences when the speaker was understood to be knowledgeable of the location.

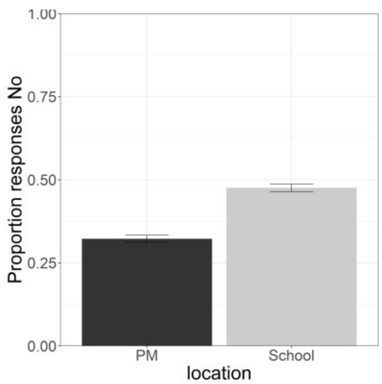


Figure 1. Proportion of inference response by location (School is familiar)

Response time. Along with measuring participants' responses, we also measured the time taken to respond to the question. Since responding "yes" is known generally to take less time than responding "no", we included Response Type as a fixed effect in the model and tested the interaction with Speaker Knowledge. If inferences are costly to compute, the lack of such computation (corresponding to the "yes" response) should yield faster reading times. As shown in Figure 2, we see the expected main effect of Response Type whereby "yes" is indeed faster than "no" (β=.172, SE =.020, t=8.540, p<.001) but no effect of Speaker Knowledge (β=.016, SE =.017, t=.893, p=.372). However, there was an interaction between Speaker Knowledge and Response Type (β=-.077, SE=.037, t=-2.077, p=.038)**,** whereby the slowest condition was for the inference ("no") for the knowledgeable speaker condition (school) and the fastest condition was for no inference ("yes") for the non-knowledgeable speaker (PM)."

Analysing simple effects showed that, based on Speaker Knowledge, there was no difference responding "no" (β=.023, SE=.028, z=.797, p=.856) but a marginal difference in responding "yes" (β=-.054, SE=.023, z=-2.404, p=.076)

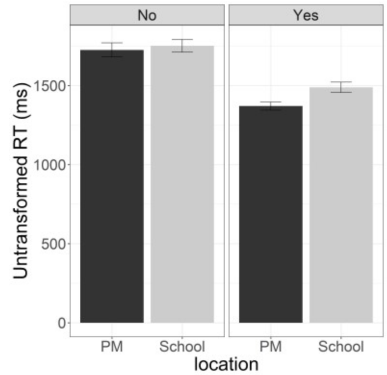


Figure 2. Response times by response and location. ( "No" signals inference)

**Reading times.** We measured the time participants spent reading the critical utterances before pressing a button to continue. We found no difference in "no" response times between conditions, thus there is no evidence that Speaker Knowledge influences the speed at which participants confirm the presence of the inference, despite seeing the different rates of inference. It is possible that *informativity-based* inferences arise automatically, before we ask for participants' judgements. If this is the case, then it is likely that we would see evidence of the cost of the inference computation at the point of the utterance, rather than downstream when asking for judgements. Thus, there should be greater reading times in the Knowledgeable Speaker condition (School). However, this prediction regarding an effect of Speaker Knowledge on utterance reading times was not borne out. Furthermore, there was no difference in reading times based on the response participants went on to give nor was there an interaction between Speaker Knowledge and Response (p's>.146)

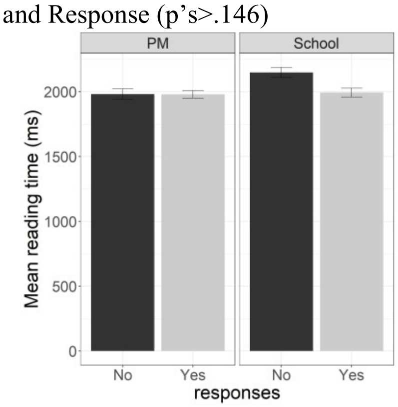


Figure 3. Mean utterance reading times by location and subsequent inference response ( "No" signals inference)

## Discussion

Experiment 1 found that *informativity-based* inferences rely on utterances being produced by knowledgeable speakers. Based on the reading time and response time data, *informativity-based* inferences do not seem to be derived automatically; rather we see evidence of inference computation when participants are prompted to answer a question about the potential inference ("Was it the same a few months ago?"). There was no difference between conditions in reading time for the sentences; however an interaction was found for response times. We demonstrated that participants took longer to respond "no" to the question "Was it the same a few months ago?" irrespective of the Speaker Knowledge manipulation. It is unclear however if this delay can be attributed to the derivation of an *informativity* inference (i.e. *No, it was not the same*) or if the increase response times is a consequence of it taking longer for participants to respond "no".

Thus, in Experiment 2 we replicate Experiment 1 with one small modification; we changed the wording of the question from "Was it the same a few months ago?" to "Was it different a few months ago?" Thus, it is now the "yes" responses which are indicative of an inference. If *informativity* inferences are costly to compute then it is expected that it will take longer for participants to respond "yes" than "no". As we will show, participants take longer to respond "no" than "yes" in Experiment 2 mirroring the findings from Experiment 1. We also compare the responses across the two experiments and find that a "yes" response in Experiment 2 (where "yes" signals an inference) took longer than a "yes" response in Experiment 1 (where "yes" did not indicate an inference)."

## Experiment 2

We recruited 213 participants Prolific (N=110) and the student population at University of Edinburgh (N=103). All participants were proficient English speakers.

### Method

The same materials as in Experiment 1 were used. The key difference between the two studies is the polarity of the question. Here in Experiment 2 we asked participants "Was it different a few months ago?" Thus, Experiment 1 and Experiment 2 differ in the yes/no answer which indicates the computation of an inference: In Experiment 2, "yes" responses now correspond to an inference response and "no" corresponds to a no-inference response.

### Results

Data analysis procedure same as in Experiment 1.

**Inference rates.** As in Experiment 1, there was a greater rate of inference responses (here, "yes" responses) in the knowledgeable speaker condition (see Fig 4). We found a main effect of Speaker Knowledge (β=.923, SE=.073, z =12.577 p<.001); participants again derived more informativity inferences in the familiar location.

**Response time.** The response time data shows a markedly different pattern to that of Experiment 1. Unlike in Experiment 1, the lack of inference computation (as indicated here in Experiment 2 via the "no" response) does not differ in response time to that of making an inference (see Fig.5). The results indicate there is an interaction between Speaker Knowledge and Response Type (β=.081, t= 2.092, SE=. 039, p = .037) however, follow up analyses did not did not indicate any significant differences between conditions.

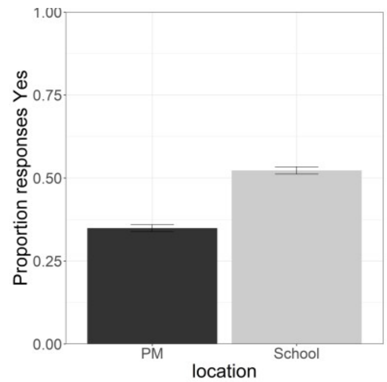


Figure 4. Proportion of inference response by location

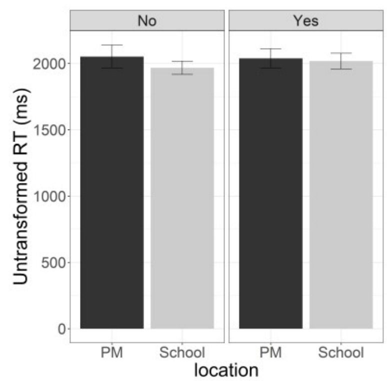


Figure 5. Response times by response and condition ("Yes" signals inference)

**Reading times.** As in Experiment 1 we find no main effect of Speaker Knowledge on reading time (β=-.020, t=-.947, SE = .021, p=.344) nor of response type (β=.015, t=.857, SE = .017, p=.392). We do see an interaction between Speaker Knowledge and Response β=-.0771, t=-2.414 SE= .032, p=.016). Follow ups indicate that this is driven by participants who responded "yes" took longer to read in the familiar vs unfamiliar location (β=-.059, SE=.034, z =-2.464 p=.066).

**Cross experiment analysis**

Of particular interest is the comparison in reaction times across experiments. In Experiment 1 "no" responses were compatible with an inference response whereas in Experiment 2 "yes" responses corresponded to the inference. The goal in comparing response times across experiments is to provide insight into the cost of deriving *informativity* inferences. We aggregated the data from Experiment 1 and 2 and compared response times as a function of Response Type and Experiment.

The analysis showed a main effect of Response Type (β=.175, t= 6.175, SE=.041, p < .001), of Experiment (β=.255, t=8.730, SE=.255, p<.001), and an interaction (β=-.193, t=-6.870, SE=.028, p<.001). Pairwise contrasts demonstrate a significant difference responding "yes" or "no" in Experiment 1 (β=.175, z=8.730 SE=.021, p<.001).

Participants were significantly slower to respond "no" in Experiment 1 than they were to respond "yes". This could be due to longer time taken to respond "no" compared to "yes" or it could reflect an increase in processing time due to computing an inference.

Of interest to the present study was the comparison between responding "no" across the two experiments and responding "yes" across the two experiments. Pairwise contrasts demonstrated no significant difference in responding "no" across the studies (β=-.061, z=-1.473, SE=.042, p=.454). This would suggest that drawing an *informativity-based* inference is not a costly computation. However, this is not the full picture as the results show a significant difference in responding "yes" across the two studies (β=-.255, z=-6.175, SE=.041, p<.001). Participants took longer to respond "yes" in Experiment 2 when compared with Experiment 1, which is consistent with *informativity-inferences* being costly to compute.

## General discussion

The studies presented here investigated informativity inferences that arise when speaker contributions fail to satisfy listener expectations for cooperative levels of information exchange. When confronted with utterances that violate listeners' expectations of informativity listeners will compute additional meaning to reconcile their expectations about a conversational partner's contribution to the discourse (Grice, 1975; Kravtchenko & Demberg, 2015; Levinson, 2000).

The results on the availability of informativity inferences showed a clear effect of speaker knowledgeability; rates of inferencing increased when the speaker was talking about a location they were familiar with (i.e. knowledgeable). In particular, listeners inferred that what was asserted as the current situation did not generally hold: When told "the library walls are blue" participants inferred that the walls used to be different. This finding supports the claim that interlocutors have expectations of informativity; if there were not this expectation, then responses should have been similar across locations.

There is conflicting evidence from the reading time data with regards to how readily informativity-based inferences arise. From Experiment 1 it would appear that listeners accept trivial utterances without additional processing costs; we found no difference in reading times as a function of speaker knowledge. Possibly because the utterances did not violate expectations of informativity or because trivial utterances are a common part of conversation. It is more likely however, that in the present studies since participants were merely observing the communicative interaction rather than being a member of that interaction, expectations for informativity are lower such that the informativity violation was not deemed an unacceptable contribution.

However, Experiment 2 paints a different picture; there is an interaction between response type and speaker knowledge on reading whereby participants who responded yes (inference) in the familiar condition spent longer reading than participants who responded no. There was no difference in reading times based on response type in the unfamiliar condition. We hesitate to make any strong conclusions on the basis of this finding for a couple of reasons. Firstly, the lack of any difference in response times seen in Experiment 1; if it were the case that informativity-based inferences are computed without prompting we would have expected to see a similar effect in Experiment 1. Secondly, the difference in reading time was only marginally significant (p=.066) and thus before making strong claims we hope to replicate the effect.

More compelling are the finding from comparing reaction times to the prompt question across the Experiments which supports the argument that *informativity-based* inferences do not arise automatically. The data demonstrates that responding "yes" takes longer to do when such a response corresponds to an inference interpretation than compared to a no-inference response. While we did not find any difference in reaction times to responding "no" across the experiments this is likely due to additional processing costs associated with responding no.

A caveat to these findings is that the overall rates of inferencing were low; participants produced inference responses at around or below 50% of the time (below chance). This could suggest that participants are not engaged in computing additional meaning. However, there are three points that suggest this is unlikely. Firstly, there is a systematic difference depending on the location, suggesting that these an informativity inference, albeit infrequent, is more available when the context supports its inference (in the speaker knowledgeable condition) and secondly, this finding has been replicated across both experiments presented above. Experiment 1 asked if the situation was "the same a few months ago" whereas Experiment 2 asked if the situation was "different a few months ago". Despite these differences, inference rates are nearly identical across the two studies. Finally, there was minimal context provided which may decrease the likelihood of drawing informativity inferences. Given the items, it is quite likely that there is a ceiling of responses since things do not typically change often. Many of the items refer to changes that are infrequent and may require a considerable undertaking (e.g. painting walls or updating carpets). Thus, participants' real world knowledge may influence their reasoning about these utterances.

## Conclusion

The present experiments tap into the types of reasoning undertaken during communication and the broader questions about how and when inferences are computed. Overall, they demonstrate that listeners have pervasive expectations of cooperativity and if conversational contributions fail to satisfy this then listeners can engage in sophisticated reasoning to reconcile the mismatch in informativeness. This extends previous work on inferencing, which typically targets specific classes of words that give rise to inferences and demonstrates that broader, systematic, goal-driven inferencing arises even in the absence of cues to pragmatic enrichment.

## Acknowledgments

This work was funded by a Leverhulme ECF awarded to Alice Rees and by a Moray Endowment Fund grant awarded to Alice Rees and Hannah Rohde.

## References

Baker, R., Gill, A., & Cassell, J. (2008). Reactive redundancy and listener comprehension in direction-giving. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue* (pp. 37-45).

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H. … & Grothendieck, G. (2011). Package 'lme4'. *Linear mixed-effects models using S4 classes. R package version*, *1*(6).

Bergen, L., & Grodner, D. J. (2012). Speaker knowledge influences the comprehension of pragmatic inferences. Journal of Experimental Psychology: Learning, Memory, and Cognition, 38(5), 1450.

Bergey, C. & Yurovsky, D. (2022, preprint). Using contrastive inferences to learn about new words and categories.

Bohn, M., Tessler, M. H., Merrick, M., & Frank, M. C. (2019). Predicting pragmatic cue integration in adults' and children's inferences about novel word meanings.

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. Cognition, 100(3), 434-463.

Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive psychology*, *19*(4), 441-472.

Carlson, G. N. (1977). *Reference to kinds in English*. Doctoral dissertation, University of Massachussetts, Amherst.

Carston, R. (2004). Relevance theory and the saying/implicating distinction. *The handbook of pragmatics*, 633-656.

Franke, M. & Jäger, G. (2016). Probabilistic pragmatics, or why Bayes' rule is probably important for pragmatics. *Zeitschriftfür Sprachwissenschaft, 35,* 3–44.

Gelman, S. A. (2004). Learning words for kinds: Generic noun phrases in acquisition. *Weaving a lexicon*, 445-484.

Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.

Goodman, N. & Frank, M. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Science, 11,* 818-829.

Grice, H.P (1975). Logic and Conversation in Cole & Morgan (Eds) *Syntax and Semantics,* Vol. 9, Academic Press, New York (1975), pp. 411-458

Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, *116*(1), 42-55.

Horowitz, A. C., & Frank, M. C. (2016). Children's pragmatic inferences as a route for learning about the world. *Child development*, *87*(3), 807-819.

Kreiss E, Degen J. 2020. Production expectations modulate contrastive inference. In Proceedings of the 42[nd] Annual Conference of the Cognitive Science Society, pp. 259–65. Austin, TX

Kampa, A., & Papafragou, A. (2020). Four-year-olds incorporate speaker knowledge into pragmatic inferences. Developmental science, 23(3), e12920.

Kravtchenko, E., and Demberg, V. (2022). "Informationally redundant utterances elicit pragmatic inferences." *Cognition* 225, 105159.

Lenth, R. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.6.0.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. MIT press.

Long, M., Moore, I., Mollica, F., & Rubio-Fernandez, P. (2021). Contrast perception as a visual heuristic in the formulation of referential expressions. *Cognition*, *217*, 104879.

Papafragou, A., Friedberg, C., & Cohen, M. L. (2018). The role of speaker knowledge in children's pragmatic inferences. Child Development, 89(5), 1642-1656.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for statistical Computing, Vienna, Austria.

Rohde, H., Futrell, R., & Lucas, C. G. (2021). What's new? A comprehension bias in favor of informativity. *Cognition*, *209*, 104491.

Rubio-Fernández, P. (2016). How redundant are redundant color adjectives? An efficiency-based analysis of color overspecification. *Frontiers in psychology*, *7*, 153.

Rubio-Fernandez, P. (2019). Overinformative speakers are cooperative: Revisiting the Gricean Maxim of Quantity. *Cognitive science*, *43*(11), e12797.

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.

Singh, R. (2019). Context, content, and the occasional costs of implicature computation. *Frontiers in Psychology*, *10*, 2214.

Sperber, D,. & Wilson, D. (1986)/95. Relevance: Communication and Cognition. Oxford: Blackwell; Cambridge, Mass.: Harvard University Press. Second edition (with postface) 1995.

Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42(4)*, 1096-1104.

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology, 44(1)*, 24-31.

van Tiel, B., van Miltenburg, E., Zevakhina, N., & Geurts, B. (2016). Scalar diversity. *Journal of semantics*, *33*(1), 137-175.

Westerbeek, H., Koolen, R., & Maes, A. (2015). Stored object knowledge and the production of referring expressions: The case of color typicality. *Frontiers in psychology*, *6*, 935.

Wilson, D. & Sperber, D. (2002). Relevance Theory. In Ward, G. & Horn, L. (eds) Handbook of Pragmatics, Oxford: Blackwell.
Wu, S. A., & Gibson, E. (2021). Word order predicts cross-linguistic differences in the production of redundant color and number modifiers. Cognitive Science, 45(1), e12934.