Informational goals, sentence structure, and comparison class inference

Michael Henry Tessler
Brain and Cognitive Sciences
MIT
tessler@mit.edu

Polina Tsvilodub
Institute of Cognitive Science
Osnabrück University
ptsvilodub@uos.de

Jesse Snedeker
Department of Psychology
Harvard University
snedeker@wjh.harvard.edu

Roger Levy
Brain and Cognitive Sciences
MIT
rplevy@mit.edu

Abstract

Understanding a gradable adjective (e.g., *big*) requires making reference to a comparison class, a set of objects or entities against which the referent is implicitly compared (e.g., big for a Great Dane), but how do listeners decide upon a comparison class? Simple models of semantic composition stipulate that the adjective combines with a noun, which necessarily becomes the comparison class (e.g., "That Great Dane is big" means big for a Great Dane). We investigate an alternative hypothesis built on the idea that the utility of a noun in an adjectival utterance can be either for reference (getting the listener to attend to the right object) or predication (describing a property of the referent). Therefore, we hypothesize that when the presence of a noun (N) can be explained away by its utility in reference (e.g., being in the subject position: "That N is big"), it is less likely to set the comparison class. Across three pre-registered experiments, we find evidence that listeners integrate the noun of a sentence with syntactic information and world knowledge to infer comparison classes, consistent with a trade-off between reference and predication.

Keywords: comparison class; adjectives; information structure; reference; predication

The meanings of linguistic expressions can change dramatically depending on the context. But determining which aspects of context are relevant for understanding a speaker's message is far from understood. This issue is brought into focus when trying to understand gradable adjectives like *big*, *tall*, or *beautiful*. The utterance "That Great Dane is big" informs the listener that the referent (a Great Dane) has a relatively large size, but relative to what the speaker thinks the Great Dane is big goes unsaid: The Great Dane could be *big for a Great Dane*, *big for a dog*, *big for a four-legged creature*, as well as an infinity of other possibilities. How do human listeners determine the comparison class when faced with multiple *a priori* reasonable options?

Simple models of semantic composition posit that when an adjective combines syntactically with a noun (N), an interpretable adjectival phrase is produced by the noun providing the comparison class (e.g., big(car) $\rightarrow$ *big for a car*, small(watch) $\rightarrow$ *small for a watch*; Kamp, 1975; Cresswell, 1976). Many arguments have been laid against such a simple mapping between the noun in the sentence and the comparison class (e.g., Bierwisch, 1989; Kennedy, 2007): A *big snowman* said of a snowman that a 4-year-old built probably means something like *big relative to snowmen that 4-year-olds can build*; a *rich Fortune-500 CEO* might not be *rich relative to other Fortune-500 CEOs*. Theoretical work on comparison classes has focused on how comparison classes are integrated into a compositional semantics and what representations might be preferred (Kennedy, 2007; Solt, 2009; Bale, 2011). Yet, little is known about how human listeners decide upon a comparison class in context.

We examine the problem from a functional perspective: what goals are speakers trying to achieve when crafting their utterance, and how might these goals influence listeners' interpretations? In order to communicate a property of a referent, a speaker must achieve two informational goals: the goal of *reference* (helping the listener identify the right target) and the goal of *predication* (communicating a property of the referent) (Reboul, 2001). In simple, subject-predicate sentences of the form "*S P*", where *S* is a referential subject noun phrase and *P* is a predicate that is asserted to hold of the subject, we posit that listeners expect that the referent will be made clear by the subject N – independent of the predicate – and that speakers aim to satisfy this expectation. [1] [pt: subject NOUN PHRASE might be unclear, since we also look at sentences where the subject has no N and posit the same; we might want to add a sentence about general expectations of reference-in-subject. maybe sth like this? For simple sentences of the form "Subject Predicate", we posit that listeners expect that the referent will be made clear by the Subject - independently of the predicate, which is asserted to hold of the subject, and that speakers...] If this expectation holds when interpreting gradable adjectives, then the influence of a speaker's choice of noun on the comparison class will depend on whether the noun appears in the subject (for the goal of reference) or the predicate (for the goal of predication; Fig. 1). If the noun appears in the predicate ("That's a big {Great Dane, dog}"), then it is natural to explain the speaker's choice of noun as non-referential, but rather as a cue to the intended comparison class. In contrast, if the noun appears in the subject ("That {Great Dane, dog} is big"), then the speaker's choice of noun can potentially be *explained away* as intending to help the listener establish reference of the subject, thus serving as a weaker cue to comparison class and allowing

[1] Of course, it is not universally true that the referent is established by the subject NP (e.g., insofar as one can infer who *he* is in the sentence "He's making those outrageous tweets again.", it is because the predicate provides a cue to the referent). We posit this relation between subject NP and reference as an expectation that listeners may hold, perhaps due to information structural reasons.

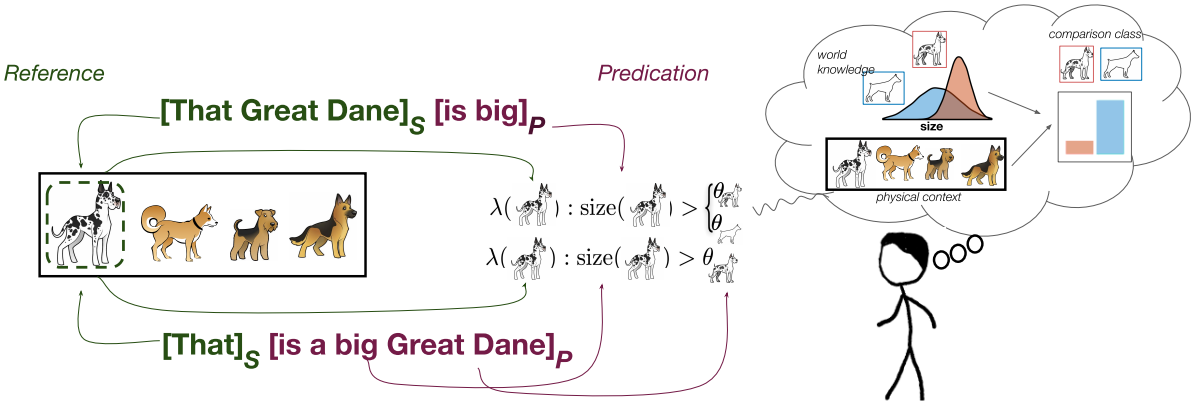

Figure 1: Cartoon of inferential account for comparison class determination. The noun (Great Dane) in a sentence can be employed either for the goal of reference (green) or predication (purple), shown in the case when this distinction is made via the syntactic position of the N (subject (S) vs. predicate (P)). When the noun is used for reference (top), a listener is left with uncertainty about what to use as the comparison class (dogs or Great Danes) and integrates their world knowledge and the physical context to make this inference. When the noun is used for predication (bottom), the listener should have less uncertainty about the comparison class: The comparison class is stipulated by the noun.

other pragmatic reasoning (e.g., world knowledge and prior expectations) to play a larger role in determining the comparison class (e.g., the Great Dane is big for a dog; Tessler, Lopez-Brau, & Goodman, 2017).

We test this reference – predication trade-off hypothesis using a syntactic manipulation wherein the N can appear either in the subject or the predicate of a sentence involving a gradable adjective (e.g., "That Great Dane is big" vs. "That's a big Great Dane"). The critical test of this manipulation is how speakers and listeners treat these sentences in the context of a referent for whom the adjective is felicitous given one comparison class but not another (e.g., *big* to describe a normal-sized Great Dane, which would be big if the comparison class is *dogs* but not *Great Danes*). We examine human judgments using three distinct dependent measures in pre-registered experiments.

Experiments

Our guiding hypothesis is that when speakers compose their utterance, the utility of a N in reference trades off with the utility of it conveying a feature value of the referent (predication) [2]; utility in reference can then "explain away" the utility of using a noun to set the comparison class. We operationalize utility in reference via the syntactic frame in which the noun phrase appears: if the noun appears in the subject of the sentence (That N is ADJ), it is likely to be used for reference and less likely to set the comparison class. If the noun appears in the predicate of the sentence (That's an ADJ N), it is unlikely to be used for reference and more likely to set the comparison class.

Table 1: Experimental items: each basic-level context had two potential targets from an either saliently small or saliently big subordinate category within the basic-level class. Items marked with * were used in Expt. 2.

Basic-level category	Smaller referent	Bigger referent
Dogs	Pug	Great Dane
Dogs	Chihuahua	Doberman
Birds	Hummingbird	Eagle
Fish	Goldfish	Swordfish
Flowers	Dandelion	Sunflower
Trees	Bonsai	Redwood
Birds*	Sparrow*	Goose*
Birds*	Canary*	Swan*
Fish*	Clownfish*	Tuna*
Flowers*	Daisy*	Peony*

In all of our experiments, we use the ADJs *big* and *small* because of the simplicity with which the feature value (i.e., size) can be conveyed through visual presentation and for which people have strong expectations about the feature value for members of different categories (e.g., Great Danes are generally big dogs; goldfish are generally small fish; etc.). Referents were always described using the size adjective consistent with these general expectations (e.g., *Great Dane* – *big*, *goldfish* – *small*), to allow for the possibility of either subordinate (Great Dane) or basic-level (dog) comparison classes. The preregistrations and full experimental procedures can be viewed at `tinyurl.com/rcsyz9f`[3].

[2] For scalar degrees, this amounts to communicating the intended comparison class of the respective gradable adjective

[3] All data and code can be found under `https://github.com/polina-tsvilodub/refpred`

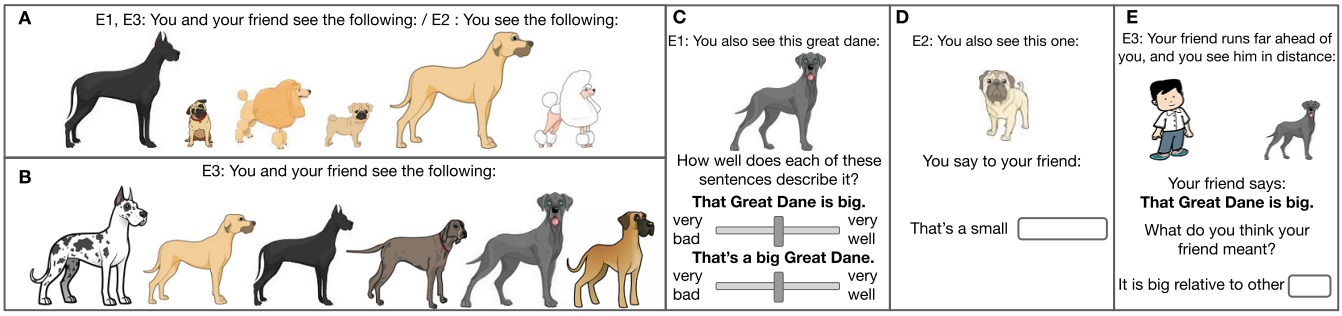

Figure 2: Overview of Experiments 1-3. A - B: Example context stimuli. A: Basic-level contexts used in Expts. 1-3. B: Subordinate context from Expt. 3. C - E: Example test questions with referents. C: Syntax Rating trial (Expt. 1) with a referent from a large-subordinate category referred to with a subordinate N. D: NP Production trial (Expt. 2) with a referent from a small-subordinate category described with a predicate-N syntactic frame. E: Comparison Class Inference trial (Expt. 3) with a referent from a large-subordinate category described with a subject-N syntactic frame using a subordinate-N label.

Experiment 1: Syntax Rating

In this experiment participants rated how well each of two sentences differing in the position of the noun described the target in context. The noun was either the basic-level (e.g., dog) or the subordinate target label (e.g., Great Dane; within-subjects).

Participants We recruited 113 participants from Amazon's Mechanical Turk; participants in all experiments were restricted to those with US IP addresses and at least a 95% work approval rating. We excluded 33 for self-reporting a native language other than English, for failing a comprehension check or providing the same responses on every trial. The experiment took about 5 minutes and participants were compensated $0.80.

Materials All experiments used the same materials. Nouns and referent pictures were chosen from five *basic-level categories* in the animal and plant domains: dogs, birds, fish, flowers, trees. Within each basic-level category, we chose target objects from *subordinate level categories* about which people have prior expectations concerning the size of members of those categories (Table 1).

Procedure Participants completed two comprehension check trials and six main trials. In the comprehension check trials, participants see a picture (e.g., a purple chair), read pairs of sentences describing it (e.g., "The chair is blue" and "The chair is yellow"), and are asked to rate on slider how well each of the sentences describes the referent.

In the main trials, participants read: "You and your friend see the following:" above a context picture with other members of the same basic-level category (e.g., a group of dogs; Figure 2A). Six different basic-level contexts were created from the five categories depicting groups of several members belonging to different subordinate categories (e.g., dogs of different breeds, including the target and filler subordinate categories, such as Great Danes, pugs and poodles; Table 1). Below the context they read "You also see this *subordinate label*" and saw the referent pictured.

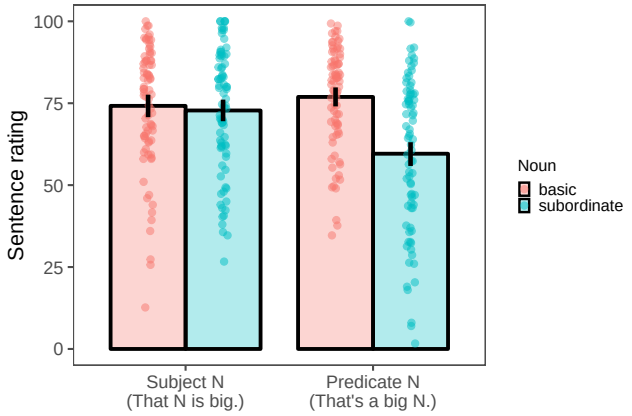

Figure 3: Experiment 1: Means and 95% bootstrapped confidence intervals (bootstrapping independent of random-effects structure) of ratings of how well the sentences described the referent when different nouns (color) appeared in different syntactic frames (x-axis). Points represent participant means within condition.

Participants rated how well two sentences described the target, using sliders ranging from *very bad* to *very well*. The sentences differed in whether the N appeared in the subject or predicate of the sentence (e.g., Predicate NP: "That's a big Great Dane"; Subject NP: "That Great Dane is big"; Fig. 2C), and the order in which the sentences and corresponding sliders appeared on the page was randomized between-subjects. Trials differed in whether the noun was the subordinate referent label (e.g., *Great Dane*) or the basic-level label (e.g., *dog*), in randomized order. Each participant saw only one of the two possible targets for each context (e.g., either the Great Dane or the pug for the dog basic-level context).

Results We found no effect of the slider presentation order (syntactic conditions), so the data was collapsed across the two conditions for all analyses. Consistent with our predic-

tion, participants substantially dispreferred sentences with the subordinate noun in predicate position compared to the subject position (Figure 3), confirmed by a Bayesian generalized linear mixed-effects model with main effects of syntax, the noun phrase, and their interaction, as well as a maximal random effects structure.[4] We found an interaction between the syntax and the NP (mean and 95% Bayesian credible interval: $\beta = -4.01[-5.84, -2.18]$), as well as an overall preference for the basic-level NPs ($\beta = -5.44[-8.09, -2.76]$) and subject-NP syntax ($\beta = -2.69[-4.77, -0.69]$). In exploratory analyses, we observed considerable variation in the by-target intercepts (e.g., *sunflower* item receives overall lower ratings), probably due to a varying basic-level label bias of the single items (the subordinate labels were more salient for some items than for others; $\beta = 9.53[5.76, 15.73]$).

Experiment 2: Free-production of noun

If the syntactic position of the N modulates the N-cue strength towards the comparison class, we would also expect speakers to produce different nouns depending on the syntactic position of the N, which we tested here.

Participants We recruited 242 participants and excluded 52 for implementation glitches, native languages other than English or failing warm-up trials more than 4 times after correction. The experiment took about 7 minutes and participants were compensated $1.00.

Procedure The main trials were divided into two blocks, and before each block, participants completed warm-up trials. The warm-up trials were designed to elicit category labels at different levels of abstraction (e.g., "Great Dane", "pug", "dog") by filling-in labeling sentences, for which they were provided corrective feedback. The same subordinate referents were used as targets in the main trials. Trial order within each warm-up and main block was randomized. We used the same contexts as in Experiment 1 and created four additional basic-level contexts (Table 1). Six contexts were randomly sampled for each participant (three per block).

On the main trials, subjects saw "You see the following:" above the context picture (as in Expt. 1; Fig 2A). Below, they read "You also see this one:" and saw the picture of the referent (e.g., a Great Dane or a pug). They were told "You say to your friend:", followed by either a subject-N or predicate-N sentence frame (between-subjects), where the noun was omitted (e.g., "That __ is big" vs. "That's a big __"; Fig. 2D). Each participant saw only either the big or the small target for each basic-level category. The free-production responses were categorized by hand into subordinate or basic-level labels of the referent. 16 uncategorizable responses (1.4%) were excluded from the analysis.

Results Participants produced basic-level nouns at a higher rate in the predicate than in the subject position (Figure 4), confirmed by a logistic Bayesian mixed-effects regression model, predicting the response category (basic-level vs. subordinate) by an intercept, the main effect of syntax and by-participant and by-referent random intercepts and a by-referent random slope effect of syntax.[5] Participants were appreciably more likely to use basic-level labels in the predicate position ($\beta = 2.25[0.74, 4.01]$).

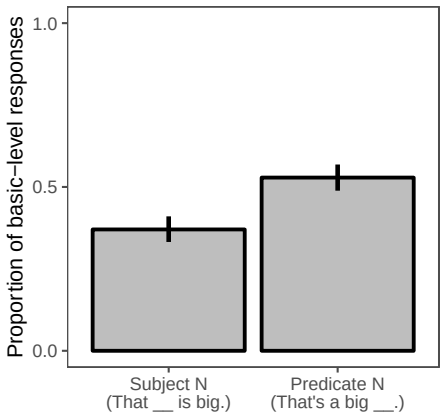

Figure 4: Experiment 2: Means and 95% bootstrapped confidence intervals of produced basic-level labels (e.g., *dog* when the referent was a Great Dane) in different syntactic frames (x-axis).

Experiment 3: Comparison Class Inference

According to our inferential account, comparison class inferences should be driven by the noun (*dog* or *Great Dane*) to the extent that the usage of the noun cannot be explained away as achieving the goal of reference. When the noun does contribute to the goal of reference, we predict comparison class inferences should be driven by the visual context.

Participants We recruited 245 participants and excluded 45 for reporting other native languages than English, failing a task comprehension check or failing warm-up trials more than 4 times after feedback. The experiment took about 9 minutes and participants were compensated $1.20.

Procedure Before the main trials, participants completed a comparison class paraphrase of the kind used in the main trials, for which they were provided corrective feedback. Following this comprehension test, participants completed two blocks of warm-up and main trials, akin to Expt. 2.

In a main trial, participants read "You and your friend see the following:" above an either subordinate-level or basic-level context picture (Fig. 2A, B). Below the context picture, they read "Your friend runs far ahead of you, and you see him in the distance" and saw a cartoon of a person standing next to the referent (e.g., a Great Dane) in the distance so that the referent size could not be judged visually (Fig. 2E). Participants read "Your friend says: [*critical sentence*]," which could vary by both syntactic position of the N (subject- vs. predicate-N) as well as the noun label. The noun label could be the subor-

[4] In lmer-style syntax: `rating ~ syntax * NP + (1 + syntax*NP | subject) + (1 + syntax*NP | target)`

[5] In lmer syntax: `response_category ~ syntax + (1 | subject) + (1 + syntax | target)`

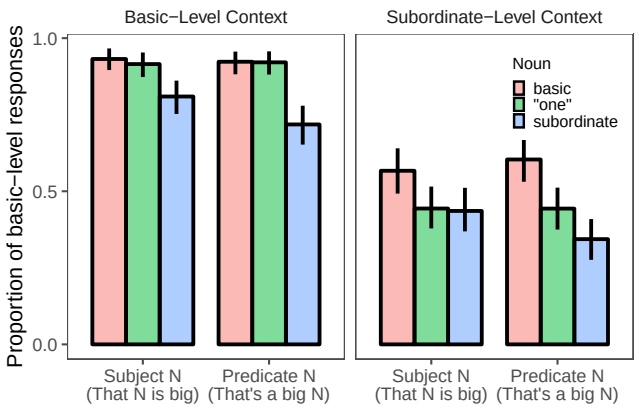

Figure 5: Experiment 3 results. Means and 95% bootstrapped confidence intervals of inferred basic-level comparison class proportions (e.g., "...big relative to other dogs").Context strongly modulated the comparison class (left vs. right panel). The noun additionally provided a cue to the comparison class (red vs. blue) bars, even in subject position. There is a hint of an interaction of the noun (red vs. blue) with the syntax.

dinate target label (e.g., Great Dane), basic level label (dog), or the underspecified noun *one* (e.g., "That one is big"). We used *one* in order to measure the baseline effect of visual context on comparison class inferences. Participants were asked "What do you think your friend meant?", to which they responded in the sentence frame: "It is {big, small} relative to other _." with their inferred comparison class (Fig. 2E).

Participants completed 12 trials, seeing exactly one trial in each condition (syntactic frame [subject vs. predicate], visual context [subordinate vs. basic], noun [subordinate vs. basic vs. *one*]), the order of which was randomized. The responses were categorized as basic-level and subordinate target labels. There were six superordinate responses which were collapsed with the basic-level responses. 39 uncategorizable responses (1.6%) were excluded from the analysis.

Results We observed substantial differences in what participants tend to decide was the most appropriate comparison class, varying by context, noun, and syntactic frame (Figure 5).

To test our pre-registered predictions, we constructed a Bayesian logistic mixed-effects regression model that predicted the response category (basic- vs. subordinate-level labels) from the syntax, context, the noun and the pair-wise two-way and three-way interactions, and a maximal random effects structure by-participant and by-referent.[6]

In addition to a global preference for basic-level comparison classes ($\beta = 2.31[1.41, 3.31]$), we observed main effects of noun-label, with basic-level labels receiving more basic-level comparison classes than the underspecified *one* ($\beta = 1.23[0.3, 2.2]$) and subordinate-level labels receiving fewer basic-level comparison classes than *one* ($\beta = 1.57[0.6, 2.6]$). Contra the simple account of comparison class determination wherein the noun in the sentence determines the comparison class, we observed a large main effect of context: more basic-level comparison classes were inferred from the basic-level than the subordinate-level context ($\beta = -2.53[-3.35, -1.90]$; Fig. 5, left vs. right facets). We see that the basic noun vs. subordinate noun difference was maintained even in the subject position (Fig. 5; red vs. blue bars), consistent with the inferential account in which the choice of noun is a cue to a speaker's conceptualization of the referent. We see some moderate evidence in support of basic vs. subordinate label x Syntax interaction (Fig. 5; red vs. blue bars vs. x-axis; 94.9% of the posterior distribution of the interaction was greater than 0, analogous to a one-tailed test).

To further explore the Noun x Syntax interaction, we built a regression model that assumed only a fixed-effect of context [7] and confirmed the NP (basic vs. sub) x Syntax interaction ($\beta = -0.49[-0.86, -0.05]$). Examining the syntax interaction in the context of NP vs. *one* contrast, we found that 95.1% of the posterior distribution of subordinate-NP vs. *one* x Syntax interaction was less than 0 ($\beta = -0.38[-0.84, 0.07]$) whereas only 62% of the posterior of the basic-NP vs. *one* X Syntax interaction was greater than 0 ($\beta = 0.62[-0.42, 0.57]$), a suggestion of a difference we return to in the discussion.

Discussion

Understanding language requires appreciating the context in which the words are uttered. Yet, speakers almost never explicitly articulate what features of context are relevant, but leave it to listeners to pragmatically reconstruct. Inferring comparison classes for relative adjectives (e.g., *big*) is a case study in this larger phenomenon of pragmatic reconstruction of context. Comparison classes are employed for understanding gradable adjectives (e.g., *big*; Kennedy, 2007) and many other linguistic expressions that convey relative meanings, including vague quantifiers (e.g., "She ate *a lot* of hot dogs"; Schöller & Franke, 2017) and generic language (e.g., "Dogs are friendly *[relative to other animals]*"; Tessler & Goodman, 2019). The basic inference we measure is that listeners are more likely to use the noun phrase in the sentence as the comparison class when the noun appears in the predicate ("That's a big Great Dane") than in the subject of the sentence ("That Great Dane is big"). We propose an information-structural reason for this inference: When the noun is in the subject of the sentence (especially when it combines with the deictic "That"), its usage can be explained away by its utility in reference, whereas a predicate-noun less strongly conveys reference and hence is more likely to be produced by a speaker aiming to convey the comparison class. Across three diverse dependent measures (Syntax Rating, Noun Production, Comparison Class Inferences), we found convergent evidence for

[6] `response_category ~ syntax*NP*context + (1 + syntax*NP*context || subject) + (1 + syntax*NP*context || target)`. We set the correlation of random effects to be 0, for computational tractability.

[7] `response_category ~ syntax*NP + context + (1 + syntax*NP || subject) + (1 + syntax*NP || item)`

such an effect.

The reference-predication trade-off hypothesis provides a starting point for an integrative account for understanding how diverse contextual circumstances and cues drive inferences about the comparison class. We argue that the utility of a noun phrase for reference can be modulated based on the syntactic position of the noun, but the syntactic distinction of subject vs. predicate is just one cue for referential vs. predicative uses. In Expt. 3, we found that comparison class inferences were driven by a subordinate noun (in comparison to the noun "one") more so when the noun appeared in the predicate of the sentence than when it appeared in the subject. We hypothesized this effect is because the referential utility of the subordinate-noun differs by syntactic position, but we note that the context must also support this inference. The referential utility of the basic-level noun was not affected by syntactic position because in neither context (basic or subordinate) was the basic-noun an informative referring expression: *dog* is both uninformative in a context of *dogs* (basic-level context) and the context of *Great Danes* (subordinate-level context). Because the basic-noun is uninformative as a referring expression, the referential-predicative trade-off view would not expect comparison class inferences to differ across syntactic positions for the basic-NP label, which is indeed what we found. Further tests of this account should experimentally manipulate the referential utility of the NP (e.g., using *dog* in the context of other animals; Graf, Degen, Hawkins, & Goodman, 2016) and confirm its impact on inferences about the comparison class.

Our subject vs. predicate noun position manipulation is perfectly confounded with whether the adjective directly syntactically modifies the noun vs. not. Direct modification could occur in the subject of the sentence: "That big Great Dane is mine". The reference-predication hypothesis we described here would predict that even in this sentence structure, the fact that Great Dane is likely to be used referentially takes some weight off its utility as a comparison class setting noun. We plan to explore this prediction in a follow-up experiment.

The reference-predication distinction we highlight in this paper, and look at through the lens of grammatical subject vs. predicate, is similar to the distinction of discourse-given vs. discourse-novel or topic vs. comment from Information Structure. Though the precise definitions of topic vs. comment are debated (e.g., Jacobs, 2001), the broad distinction is that *topic* is what is being talked about or what is given and *comment* is what is being said of the topic or what is new (Lambrecht, 1996; Krifka, 2008). We believe this distinction is dissociable from that of reference vs. predication that we focus on in this paper (Reboul, 2001). Consider, for example, the sentence: "What's big is that Great Dane". The sentence seems appropriate in a context where the topic—what is given—is that something is big, and the comment—what's new—is "that Great Dane". Yet, "Great Dane" also seems to be establishing reference, and additionally striking, it is doing so from the predicate of the sentence. Thus, though we examine reference vs. predication through the subject-predicate distinction, we believe the communicative goals are primary in driving inferences about the comparison class and are distinct from the topic-comment distinction from Information Structure.

Understanding comparison classes is a basic cognitive skill for interpreting a simple class of context-sensitive expressions: scalar adjectives. Very soon after children start producing their first scalar adjective—*big*—they seem to understand its context sensitive behavior and can flexibly switch between contexts (e.g., that a mitten, which is a small mitten, might also be *big for a doll*; Ebeling & Gelman, 1994). The kinds of cues that have been shown to modulate comparison class inferences in young children have been rather dramatic cues (e.g., "is the mitten big for the doll?"), though 2-year-olds appear sensitive to the specificity of the noun alone when interpreting adjectives (Mintz & Gleitman, 2002). The problem that the language learner faces goes beyond inferring the comparison class in the moment: Young children are jointly learning the meaning of the nouns and adjectives along with trying to construct the appropriate comparison classes to interpret the utterances they hear. Understanding children's sensitivity to the cues we investigate here can provide some hints as to how they are able to accomplish the incredible feat of learning language.

References

Bale, A. C. (2011). Scales and comparison classes. *Natural Language Semantics*, *19*, 169–190.

Bierwisch, M. (1989). The semantics of gradation. *Dimensional adjectives*, *71*(261), 35.

Cresswell, M. J. (1976). The semantics of degree. In *Montague grammar* (pp. 261–292). Elsevier.

Ebeling, K. S., & Gelman, S. A. (1994). Children's use of context in interpreting "big" and "little". *Child Development*, *65*(4), 1178–1192.

Graf, C., Degen, J., Hawkins, R. X., & Goodman, N. D. (2016). Animal, dog, or dalmatian? level of abstraction in nominal referring expressions. In *38th annual meeting of the cognitive science society.*

Jacobs, J. (2001). The dimensions of topic-comment. *Linguistics*, *39*(4; ISSU 374), 641–682.

Kamp, J. A. W. (1975). Two theories about adjectives. In E. L. Keenan (Ed.), *Formal semantics of natural language.* Cambridge University Press, Cambridge, England.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and philosophy*, *30*(1), 1–45.

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, *55*(3-4), 243–276.

Lambrecht, K. (1996). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents* (Vol. 71). Cambridge university press.

Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: The incremental and restricted nature of

early adjective acquisition. *Cognition*, *84*(3), 267–293. doi: 10.1016/S0010-0277(02)00047-1

Reboul, A. (2001). Foundations of reference and predication. In M. Haspelmath (Ed.), *Language typology and language universals. an international handbook, vol.1*. Walter de Gruyter.

Schöller, A., & Franke, M. (2017). Semantic values as latent parameters: Testing a fixed threshold hypothesis for cardinal readings of few & many. *Linguistics Vanguard*, *3*(1).

Solt, S. (2009). Notes on the Comparison Class. In *International workshop on vagueness in communication.*

Tessler, M. H., & Goodman, N. D. (2019). The language of generalization. *Psychological Review*, *126*(3), 395.

Tessler, M. H., Lopez-Brau, M., & Goodman, N. D. (2017). Warm (for winter): Comparison class understanding in vague language. In *39th annual meeting of the cognitive science society.*