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 P. 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 use the noun as a cue to infer comparison classes consistent with a trade-off between reference and predication. This work highlights the complexity of the relation between the form of an utterance and its meaning.

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

Introduction

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 large 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, an interpretable adjectival phrase is produced by using the noun as the comparison class (e.g., big(Great Dane) $\rightarrow$ *big for a Great Dane*, small(goldfish) $\rightarrow$ *small for a goldfish*; Kamp, 1975; Cresswell, 1976). There are intuitive reasons to doubt that such a simple mapping between the modified noun and the comparison class will work in general (e.g., a *rich Fortune-500 CEO* might not be rich relative to other *Fortune-500 CEOs*; Bierwisch, 1989; Kennedy, 2007), but research on comparison classes has eschewed the question of how to determine the comparison class, instead focusing on representational issues about how to integrate a comparison class (once determined) into a compositional semantics (Kennedy, 2007; Solt, 2009; Bale, 2011). The simple syntactic account could be generalized into one in which non-modified nouns in the sentence could be used as the comparison class (e.g., "That Great Dane is big" $\rightarrow$ *big for a Great Dane*). These syntactic mechanisms, however, would have nothing to say about the role that world knowledge or the physical environment might play in influencing comparison classes.

We consider the problem of comparison class inference 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: *reference* (identifying the right target object) and *predication* (ascribing a property to the referent) (Reboul, 2001). In simple sentences of the form "*Subject Predicate*", we posit that listeners expect reference to be established by the *Subject* (independent of the *Predicate* asserted to hold of the subject) and that speakers aim to satisfy this expectation.[1] From this perspective, the noun in a sentence is a cue to the comparison class (Fig. 1): If the noun appears in the predicate ("That's a big Great Dane"), the speaker's noun choice is likely non-referential, and rather a cue to the intended comparison class. In contrast, if the noun appears in the subject ("That Great Dane is big"), then the speaker's choice of noun can be *explained away* as intending to help the listener establish reference of the subject; the noun would then serve as a weaker cue to the comparison class and allow for other pragmatic reasoning (e.g., world knowledge and perceptual cues) to play a more substantial 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 noun appears either in the subject or the predicate of a sentence involving a grad-

[1] Of course, it cannot always be taken for granted that the referent is established by the subject noun (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 noun and reference as an expectation that listeners may hold, perhaps due to information structural reasons.

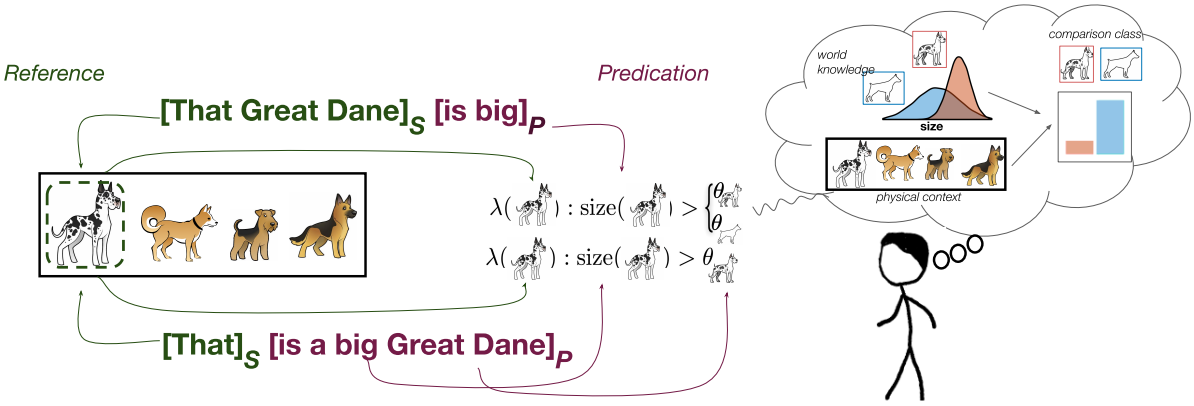

Figure 1: Cartoon of the inferential account for comparison class determination. The noun (Great Dane) in a sentence can be employed either for the goal of reference (green) or predication (purple), shown in the case when this distinction is made via the syntactic position of the noun (subject S vs. 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.

able adjective (e.g., "That Great Dane is big" vs. "That's a big Great Dane"). The critical test 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 noun in reference trades-off with the utility of the noun 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 NOUN 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 NOUN), it is unlikely to be used for reference and more likely to set the comparison class.

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 because they convey a salient feature about which participants likely have strong expectations for 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]

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*

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. 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 re-

[2] For scalar degrees, a noun conveying a feature value amounts to setting the comparison class of the respective gradable adjective

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

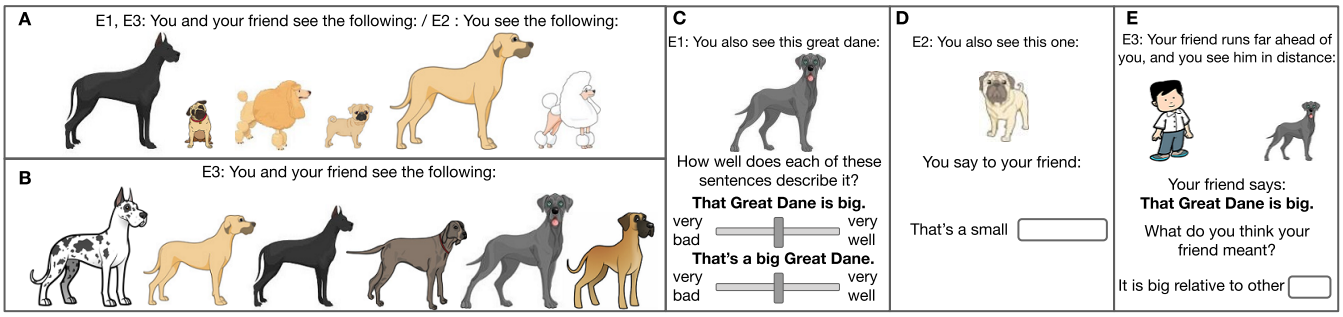

Figure 2: Overview of Experiments 1-3. A - B: Example context stimuli. A: Basic-level contexts used in Expts. 1-3. B: Subordinate context from Expt. 3. C - E: Example test questions with referents. C: Syntax Rating trial (Expt. 1) with a referent from a large-subordinate category referred to with a subordinate-level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

stricted to those with US IP addresses and at least a 95% work approval rating. We excluded 33 for reporting a native language other than English, 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 saw 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 were asked to rate on a slider how well each of the sentences described 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.

Participants rated how well two sentences described the target, using sliders ranging from *very bad* to *very well*. The sentences differed in whether the noun N appeared in the subject or predicate of the sentence (e.g., Predicate N: "That's a big Great Dane"; Subject N: "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 prediction, 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 noun-label (mean and 95% Bayesian credible interval: $\beta = -4.01[-5.84, -2.18]$), as well as an overall preference for the basic-level nouns ($\beta = 5.44[2.76, 8.09]$) and subject-N syntax ($\beta = 2.69[0.69, 4.77]$). In exploratory analyses, we observed considerable variation in the by-target intercepts (e.g., *sunflower* item received 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 noun modulates the strength of the cue the noun provides towards the comparison class, we would also expect speakers to produce different nouns depending on the noun's syntactic position, which we test here.

Participants We recruited 242 participants and excluded 52 for implementation glitches, non-native English language, or failing warm-up trials more than 4 times. The experiment

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

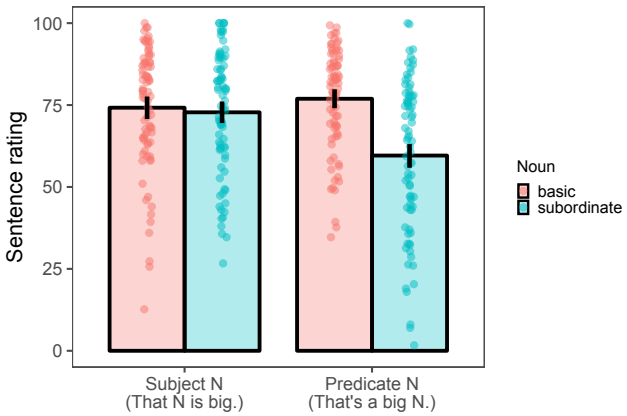

Figure 3: Experiment 1 mean ratings for how well sentences which differed in the syntactic position of the noun (x-axis) and the noun-label (color) described the referent, a typically-sized member of the subordinate category (e.g., a normal-sized Great Dane). Points represent participant means within condition. Error-bars denote bootstrapped 95% confidence intervals (bootstrapping independent of random-effects structure).

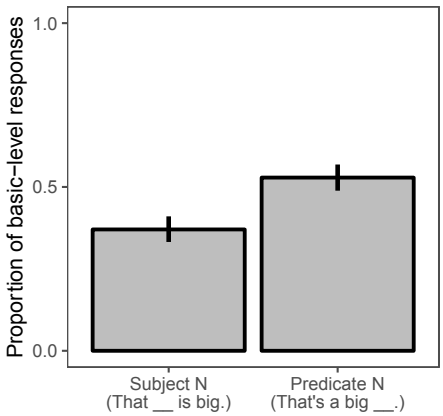

Figure 4: Experiment 2: Proportions of freely-produced basic-level labels (e.g., *dog*) in different syntactic frames (x-axis) when the referent was a typically-sized member of a subordinate category (e.g., a normal-sized Great Dane). Error-bars denote 95% bootstrapped confidence intervals.

took about 7 min 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 ($\beta = 2.25[0.74, 4.01]$).

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. Our first two experiments support this view: Participants dispreferred sentences like "That's a big Great Dane" when the subordinate comparison class was infelicitous for the referent (i.e., the Great Dane was not big for a Great Dane). In this experiment, we provide a more direct test of our account by explicitly measuring comparison class inferences.

Our experimental design manipulates three factors within-subject: syntactic position of the noun, level-of-abstractness of the noun (basic-level vs. subordinate-level vs. underspecified; e.g., *dog* vs. *Great Dane* vs. *one*), and visual context (e.g., other dogs vs. other Great Danes). Foremost, the inferential account provides a natural avenue for visual context to influence comparison class inferences: The noun in the sentence may not provide the comparison class if the visual context is very salient. Second, when the noun's usage can be explained away by its utility in reference, comparison class inferences should be more strongly driven by world knowledge (e.g., Great Danes are big dogs) or the visual context. We include an underspecified noun condition using the anaphoric "one" to provide a base-line measure of the influence of visual context on comparison class inferences: In a context with varying types (basic-level context; Fig. 2A), anaphoric "one" should be interpreted as "dog"; when the context provides animals of the same type (subordinate-level context; Fig. 2B),

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

anaphoric "one" is more likely to be interpreted more narrowly ("Great Dane"; Goldberg & Michaelis, 2017).

Participants We recruited 245 participants and excluded 45 for either 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 the main trials, participants read "You and your friend see the following:" above a context picture of either subordinate-level or basic-level distractors (Fig. 2A, B). Below the context picture, they read "Your friend runs far ahead of you, and you see him in the distance" with a cartoon person standing next to the referent (e.g., a Great Dane) in the distance so that the referent size could not serve as a cue to the comparison class (Fig. 2E). Participants read "Your friend says: [critical sentence]", which could vary by both syntactic position of the noun and the noun-label's level-of-abstractness (e.g., "That Great Dane is big", "That's a big dog", "That one is big", etc.). Participants were asked: "What do you think your friend meant?" and responded in the sentence frame: "It is big (small) relative to other __" (Fig. 2E). Participants completed 12 trials which could vary by syntactic frame [subject vs. predicate], visual context [subordinate vs. basic], and noun [subordinate vs. basic vs. *one*].[6]

Results Responses were categorized by hand as either basic-level or subordinate labels. Six of participants' responses were superordinate category labels (e.g., "animals"), which we collapsed with the basic-level responses. 39 uncategorizable responses (1.6%) were excluded from the analysis.

To test our predictions, we constructed a Bayesian logistic mixed-effects regression model that predicted the response category (basic- vs. subordinate-level comparison class) from the syntax, context, noun-label and the pair-wise two-way and three-way interactions, with a maximal random effects structure afforded by our design (Barr, Levy, Scheepers, & Tily, 2013).[7] A simple, syntactic account of comparison class determination would hold that the noun in the sentence determines the comparison class: the same sentence should receive the same comparison class regardless of context. Contra this account, we observe a large main effect of context: given the exact same sentence, more basic-level comparison classes were inferred overall from the basic- than subordinate-level

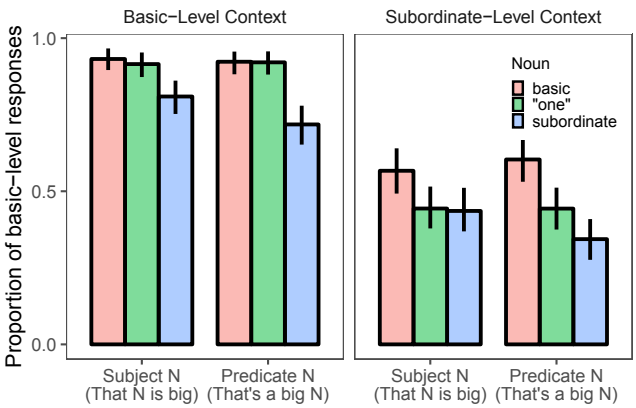

Figure 5: Experiment 3 results. Proportions of inferred comparison classes in terms of basic-level responses (e.g., "...big relative to other dogs") as a function of syntactic position of the noun (x-axis), noun-label (color), and context (facets). Context strongly modulated the comparison class (left vs. right panel). The noun additionally provided a cue to the comparison class (red vs. blue) bars, even in subject position. The effect of noun (red vs. blue) is modulated by syntax. Error-bars denote bootstrapped 95% confidence intervals.

context ($\beta = 1.88[1.49, 2.31]$; Fig. 5, left vs. right facets) and in particular, for our baseline anaphoric *one* condition ($\beta = 0.37[0.10, 0.64]$). We additionally observe a main effect of noun-label regardless of the syntactic position of the noun, arguing against an account wherein only syntactically-modified nouns ("ADJ NOUN") provide the comparison class: basic-level nouns were overall more likely to trigger basic-level comparison classes compared to subordinate nouns ($\beta = 2.01[1.37, 2.71]$). These noun-labels influenced comparison classes above and beyond the baseline for the perceptual contexts, as indexed by the anaphoric *one* condition: basic vs. *one* ($\beta = 0.60[-0.47, 1.70]$) and subordinate vs. *one* ($\beta = -1.40[-2.17, -0.66]$). We also observe that a subordinate noun can be the minority comparison class response even when the noun is syntactically modified by the adjective (e.g., "big Great Dane" $\rightarrow$ for a dog; Fig. 5, left-facet, blue-bar).

We find evidence in support of the Noun (basic vs. sub) x Syntax interaction predicted by the reference – predication trade-off hypothesis: Participants provided more subordinate-level comparison classes when the subordinate-level noun appeared in the predicate than when it appeared in the subject, in comparison to the basic-level noun (red vs. blue bars X x-axis; $\beta = 0.47[0.02, 0.95]$). We further examine the syntax X noun interaction in the context of the N vs. *one* contrasts and find suggestive evidence that this interaction is driven by the subordinate noun.Specifically, we find a 90.0% probability that the subordinate-N vs. *one* x Syntax interaction term was less than 0 (i.e., more subordinate comparison classes when the noun was in the predicate; $\beta = -0.36[-0.93, 0.22]$) in contrast to only a 65.5% probability of the basic-N vs. *one* X Syntax interaction being greater than 0 (i.e., more basic-

[6] Due to a coding error, condition balancing occurred at the level of individual factors independently but not jointly (i.e., participants did not complete all 12 unique trial types, but did complete equal numbers of each level of each factor).

[7] `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.

level comparison classes when the noun was in the predicate; $\beta = 0.11[-0.44, 0.68]$).[8] These results are consistent with listeners entertaining a trade-off between the communicative goals of reference and predication when reasoning about the comparison class for an adjectival utterance.

Discussion

Understanding language requires appreciating the context in which the words are uttered. Yet, speakers almost never articulate explicitly what features of context are relevant and 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. In addition, comparison classes play a role in understanding many kinds of 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). Additionally, studying comparison classes particularly stresses the complexity of the relation between the form of an utterance and its meaning.

The basic inference we highlight 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, its usage can be explained away by its utility in reference (especially when it combines with the deictic "That"), whereas a predicate-noun less strongly conveys reference and hence is more likely to be produced by a speaker aiming to use the noun to convey 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 more so when the noun appeared in the predicate of the sentence than when it appeared in the subject. We hypothesized this effect is due to the referential utility of the subordinate-noun differing 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 the noun's syntactic position because in neither context (basic or subordinate) was the basic-noun an informative referring expression: *dog* is both referentially uninformative in a context of *dogs* (basic-level context) and the context of *Great Danes* (subordinate-level context). As a result, the referential – predicative trade-off view would not expect comparison class inferences to differ across syntactic positions for the basic-level noun, which is indeed what we found. At the same time, in the subordinate-level context, we do observe a higher rate of basic-level comparison class inferences for basic-level nouns, regardless of the syntax: Since the context makes the basic-level noun referentially uninformative, listeners may reason that the usage of the noun was to set the comparison class. Further tests of this account should experimentally manipulate the referential utility of the noun (e.g., "dog" in the context of other animals; Graf, Degen, Hawkins, & Goodman, 2016) and confirm its impact on inferences about the comparison class.

In our experimental paradigm, the subject vs. predicate noun position manipulation is perfectly confounded with whether or not the adjective syntactically modifies the noun. Direct modification can occur in the subject of the sentence: "That big Great Dane is my favorite". Under the assumption that the subject is expected to establish reference, the reference – predication hypothesis here would also predict that the noun "Great Dane" is unlikely to set the comparison class. We plan to explore this prediction in a follow-up experiment.

The reference – predication distinction we highlight in this paper is similar to the distinction of 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 and *comment* is what is being said of the topic (Lambrecht, 1996; Krifka, 2008). We believe this distinction is dissociable from that of reference vs. predication (Reboul, 2001). Consider, for example, the sentence: "What's big is that Great Dane". The sentence seems appropriate in a context where the topic is that something is big and the comment is that it is "that Great Dane". Yet, "Great Dane" also seems to be establishing reference and, additionally striking, it is doing so from the grammatical predicate of the sentence.[9]

Understanding comparison classes is a basic cognitive skill used for interpreting a simple class of context-sensitive expressions: scalar adjectives. Very soon after children start producing their first scalar adjective (i.e., *big*), they seem to understand its context-sensitive behavior and flexibly switch between contexts (e.g., that a small mitten might also be *big for a doll*; Ebeling & Gelman, 1994). The kinds of cues 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.

[8] These probabilities are computed by examining the proportion of the posterior distribution over the parameter than is above or less than 0. This comparison is analogous to "1-tailed" statistical tests from the frequentist tradition.

[9] Though we examine reference vs. predication through the grammatical subject – predicate distinction, we believe the communicative goals, not grammatical positions, are primary in driving inferences about the comparison class.

Acknowledgments

This material is based upon work supported by the National Science Foundation SBE Postdoctoral Research Fellowship Grant No. 1911790 awarded to M.H.T. as well as support from Elemental Cognition and a Google Faculty Research Award to R.P.L.

References

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

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. journal of memory and language, 68(3), 10.1016/j.jml.2012.11.001. https://doi.org/10.1016/j.jml.2012.11.001. *Journal of Memory and Language*, *68*(3).

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.

Goldberg, A. E., & Michaelis, L. A. (2017). One among many: Anaphoric one and its relationship with numeral one. *Cognitive Science*, *41*, 233–258.

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.*