Using contrastive inferences to learn about new words and categories

Claire Bergey and Daniel Yurovsky

S1 Experiment 1

In addition to the analyses reported in the main text, we ran a pre-registered linear mixed effects model predicting target choice from the presence of an adjective in the utterance, the adjective type (size or color), and the display type (unique target display or contrastive display) (Table S1). People were more likely to choose the target if there was an adjective in the utterance ($\beta_{adjective}$ = 2.21, t = 7.18, $p = < .001$), and were more overall likely to choose the target on unique target trials (β_{unique} = 3.81, t = 10.68, $p = < .001$). There was an interaction between the presence of an adjective and the type of adjective, such that people were especially likely to choose the target when there was a size adjective in the utterance ($\beta_{adjective*size}$ = 0.95, t = 2.18, p = .029). There was a three-way interaction between the presence of an adjective, the type of adjective, and the search type such that the contrastive strength of size over color was stronger in the contrastive trials than the unique target trials ($\beta_{adjective*size*unique}$ = -3.06, t = -2.61, p = .009).

Table S1: Full model of target choice from Experiment 1. Model specification is chose_target ∼ utterance_type * adjective_type * display_type + (1 + utterance_type | subject).

term	estimate	z-value	p-value	d [95% CI]
intercept	-2.07	-7.94	< .001	-1.14 [-1.42, -0.86]
utterance type: adjective (vs. no adjective)	2.21	7.18	< .001	1.22 [0.88, 1.55]
adjective type: size (vs. color)	-0.17	-0.46	.646	-0.09 [-0.48, 0.3]
display type: unique target (vs. contrastive)	3.81	10.68	< .001	2.1 [1.72, 2.49]
adjective * size	0.95	2.18	.029	0.53 [0.05, 1]
adjective * unique target	1.32	1.22	.223	0.73 [-0.44, 1.91]
size * unique target	-0.17	-0.37	.709	-0.1 [-0.6, 0.41]
adjective * size * unique target	-3.06	-2.61	.009	-1.69 [-2.96, -0.42]

Figure S1 shows referent choice in both the unique target display trials and the contrastive display trials. Unique target displays had one unique referent (the target) and two identical distractors that differed from it both in shape and the critical feature. Contrastive displays had a target, a contrastive pair which matched the target in shape but had a different critical feature, and a lure which matched the target on the critical shape but differed from it on the critical feature.

S1.1 Modeling Experiment 1 with continuous semantics

Degen, Hawkins, Graf, Kreiss, & Goodman (2020) capture asymmetries in description of size and color by positing that different features have different semantic strength. They posit that color has stronger semantics than size, such that "red table" is a better literal description of a small red table

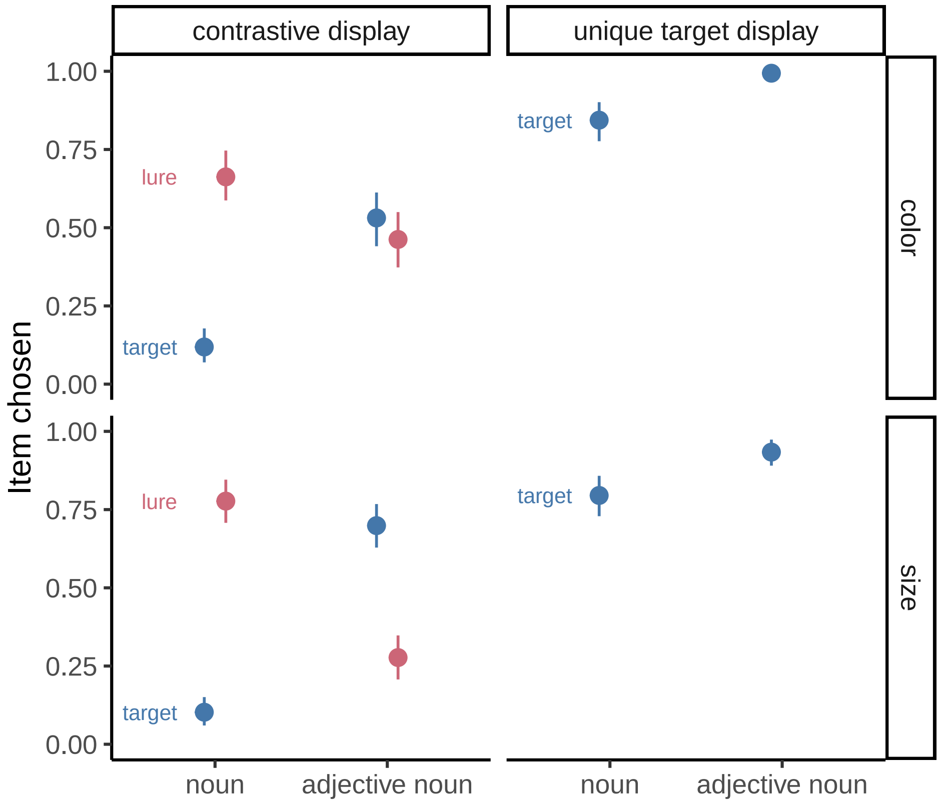

Figure S1: Referent choice in both the contrastive display trials and the unique target display trials.

than "small table" is. Under these assumptions, RSA using these continuous semantics explains people's tendency to mention color more often than size in a variety of tasks. Can their model explain the asymmetry we find between color and size in Experiment 1?

In Experiment 1, we found that people more consistently choose the target using contrastive inferences about size than color. We incorporated their continuous semantics into our RSA model of referent choice, which reasons over possible lexicons. In Figure S2 we show the difference in referent choice when a feature has low semantic strength (0.8) compared to high semantic strength (0.99). A feature with low semantic strength results in a weaker contrastive inference (reduced choice of the target in the *adjective noun* trials) compared to a feature with high semantic strength. Degen et al. (2020) find that color has stronger semantics than size, which would result in a stronger contrastive inference about referent choice when color adjectives are used. This is not what we find: people make stronger contrastive inferences about referent choice when *size* adjectives are used. Thus, while a model with continuous semantics could in principle explain the asymmetry we find, it would need to have stronger semantic values for size than color. We note that while the same continuous semantics do not explain both our data and the production data from Degen et al. (2020), neither does the model we propose explain the production data. We leave it to future work to form a more complete account of color-size asymmetries in both production and comprehension.

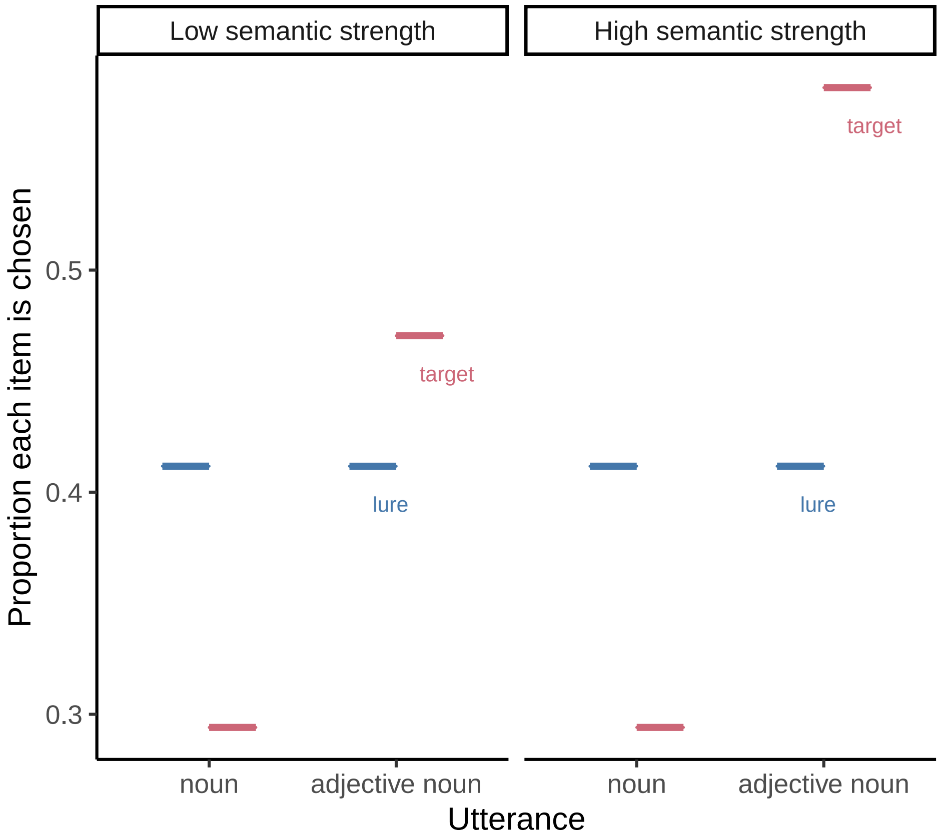

Figure S2: Results of modeling target choice in Experiment 1 using continuous semantics. Stronger continuous semantics predict higher choice of the target, while weaker continuous semantics predict lower choice of the target.

S2 Experiment 2

S2.1 Experiment 2 Prevalence Judgments

The full regression of prevalence judgments, also reported in the main text, is in Table S2.

Table S2: Full model of prevalence judgments from Experiment 2. Model specification is percentage ~ adjective_
utterance_type * context_type + (utterance_type | subject).

term	estimate	z-value	p-value	d [95% CI]
intercept	52.16	22.40	< .001	2.18 [2.13, 2.2263771652
adjective type: size (vs. color)	4.73	1.46	.146	0.86 [NaN, 1.3268029859
utterance type: adjective (vs. no adjective)	-10.22	-3.37	.001	NaN [NaN, NaN]
context: within-category contrast display (vs. between-category contrast)	3.92	1.63	.104	0.75 [NaN, 1.1894964325
context: same feature display (vs. between-category contrast)	-1.48	-0.62	.537	NaN [NaN, 0.6457544494
size * adjective	0.04	0.01	.993	-1.8 [NaN, 1.1680352382
size * within-category contrast display	-1.37	-0.41	.684	NaN [NaN, 0.911570725
size * same feature display	-0.60	-0.18	.859	NaN [NaN, 0.986648991
adjective * within-category contrast display	-1.58	-0.46	.644	NaN [NaN, 0.898975014
adjective * same feature display	2.13	0.63	.532	0.42 [NaN, 1.1996435106
size * adjective * within-category contrast display	-1.39	-0.29	.770	NaN [NaN, 1.141911593
size * adjective * same feature display	-1.59	-0.33	.739	NaN [NaN, 1.127990400

S3 Experiment 3

The full regression predicting Experiment 3 prevalence judgments, also reported in the main text, is shown in Table S3. The regression predicting Experiment 3 prevalence judgments among only adjective utterances and no adjective utterances (excluding alien utterance trials), also reported in the main text, is shown in Table S4.

In addition to the regressions reported in the manuscript, we two pre-registered, targeted regressions to test the effect of utterance type to more specifically in case these effects were unclear in the maximal models. First, we filtered to adjective and no adjective trials and fit a linear mixed effects model predicting prevalence judgment by utterance type with a random slope of utterance type by subject (Table S5). Participants' prevalence judgments were significantly lower when an adjective was used in the utterance (β = -9.17, t = -7.09, p = < .001). Second, we included all trials in a linear mixed effects model predicting prevalence judgment by utterance type with a random slope of utterance type by subject (Table S6). Utterances without an adjective resulted in significantly higher prevalence judgments than alien utterances (β = 7.76, t = 4.91, p = < .001), and utterances with an adjective did not result in significantly different prevalence judgments than alien utterances (β = -1.42, t = -0.91, p = .363).

Table S3: Regression predicting prevalence judgments from utterance type, context type, and adjective type in Experiment 3. Model specification is percentage ~ utterance_type * context_type * adjective_type + (utterance_type | subject).

term	estimate	statistic	p-value
intercept	47.77	20.73	< .001
utterance type: no adjective utterance (vs. alien utterance)	7.48	2.80	.005
utterance type: adjective (vs. alien utterance)	-0.64	-0.24	.808
context type: within-category contrast (vs. between-category)	-2.70	-1.23	.220
adjective type: size (vs. color)	4.44	1.33	.185
no adjective utterance * within-category contrast display	5.57	1.79	.073
adjective utterance * within-category contrast display	5.77	1.86	.064
no adjective utterance * size	-5.09	-1.32	.189
adjective utterance * size	-6.56	-1.72	.086
within-category contrast display	1.24	0.39	.696
no adjective utterance * within-category contrast display * size	-0.32	-0.07	.944
	-2.21	-0.49	.623

Table S4: Regression predicting prevalence judgments from utterance type, context type, and adjective type only among adjective and no adjective utterances (excluding alien utterances) in Experiment 3. Model specification is percentage ~ utterance_type * context_type * adjective_type + (utterance_type | subject).

term	estimate	statistic	p-value
intercept	55.24	23.76	< .001
utterance type: adjective (vs. no adjective)	-8.12	-3.46	.001
context tyep: within-category contrast (vs. between-category)	2.87	1.34	.180
adjective type: size (vs. color)	-0.66	-0.20	.845
adjective utterance * within-category contrast display	0.19	0.06	.949
adjective utterance * size	-1.47	-0.43	.665
within-category contrast display	0.92	0.30	.766
no adjective utterance * within-category contrast display * size	-1.90	-0.43	.665

Table S5: Regression predicting prevalence judgments from presence of an adjective in the utterance (excluding alien language utterances) in Experiment 3. Model specification is percentage ~ utterance_type + (utterance_type | subject).

term	estimate	std.error	statistic	df	p-value
intercept	56.59	1.49	38.00	196.00	< .001
adjective utterance (vs. no adjective utterance)	-9.17	1.29	-7.09	196.00	< .001

Table S6: Regression predicting prevalence judgments from utterance type in Experiment 3. Model specification is percentage ~ utterance_type + (utterance_type | subject).

term	estimate	std.error	statistic	df	p-value
intercept	48.83	1.47	33.18	196.00	< .001
no adjective utterance (vs. alien utterance)	7.76	1.58	4.91	196.00	< .001
adjective utterance (vs. alien utterance)	-1.42	1.55	-0.91	196.00	.363

References

Degen, J., Hawkins, R. D., Graf, C., Kreiss, E., & Goodman, N. D. (2020). When redundancy is useful: A Bayesian approach to "overinformative" referring expressions. *Psychological Review*, *127*, 591–621.