Supporting information: Comparing subjectivity with alternative accounts of adjective order

In the main text, we have described the success of adjective subjectivity in predicting ordering preferences for multi-adjective strings. In this supplement, we attempt to compare the predictions of adjective subjectivity with those of alternative proposals. Direct quantitative comparisons prove difficult because the investigations to date have been largely qualitative. While many authors have made useful theoretical contributions, most have not attempted to operationalize their hypotheses, let alone test them in large-scale behavioral studies. Indeed, most authors report only their own intuitions or the intuitions of a handful of informants, using a small set of adjectives and nouns.[1]

In order to facilitate comparison with the predictions of subjectivity, we attempt here to operationalize several previous accounts. We begin at the beginning, with the "inherentness" hypothesis from Sweet (1898) and Whorf (1945), which we operationalize by using the design of an experiment from Martin (1969). Then we look at subsectivity, a binary theoretical construct that some linguists have argued underlies ordering preferences (e.g., Truswell, 2009). Finally, we try once more to find noun-specific ordering preferences that would support an account based on concept formation, and even attempt to operationalize a concept-formability metric (McNally and Boleda, 2004; Bouchard, 2005; Svenonius, 2008). To preview the findings, we continue to find that subjectivity predicts ordering preferences with remarkable accuracy, while alternative approaches do significantly less well.

1 The predictive power of inherentness

Early accounts of adjective ordering proposed that the primary factor derterming order was the "inherentness" of the adjectives. We find perhaps the most eloquent formulation of this claim in the following quote from (Whorf, 1945, p. 5):

> English adjectives form two main cryptotypes with sub-classes. A group referring to 'inherent' qualities—including color, material, physical state (solid, liquid,

[1] Martin (1969) is the exception, reporting behavioral measures of ordering preferences and four different aspects of adjective meaning that he hypothesized would predict these ordering preferences. The best-performing measure was adjective "definiteness;" it accounts for between 32% and 55% of the variance in his preference data. (He tested two different sets of twenty adjectives. Adjective definiteness accounted for 55% of the variation in the ordering preferences for first set of adjectives; for second set, definiteness accounted for 32% of the variation.) We have seen that subjectivity seems to fare better, and in addition our study uses more adjectives, more nouns, and two different operationalizations of subjectivity.

> porous, hard, etc.), provenience, breed, nationality, function, use—has the reactance of being placed nearer the noun than the other group, which we may call one of non-inherent qualities, though it is rather the residuum outside the first group—including adjectives of size, shape, position, evaluation (ethical, esthetic, or economic). These come before the inherent group, e.g. *large red house* (not *red large house*), *steep rocky hill*, *nice smooth floor*.

Unfortunately, most authors are content with merely asserting the claim concerning inherentness, not rigorously testing it. The single outlier is Martin (1969); he documented that inherentness accounted for between 3% and 14% of the variation in the ordering preferences that he measured for 40 adjectives. Still, to allow for a direct comparison of the predictions of subjectivity with those of inherentness, it was necessary to measure their values for the same set of materials. Therefore, we ran a version of Martin's original task using the adjectives and nouns from Expt. 1, and included a replication of our subjectivity experiment.

Participants. We recruited 72 participants through Amazon.com's Mechanical Turk. Participants were compensated for their participation.

Design and methods. Participants were randomly assigned to one of two conditions: *inherentness* (n=41) or *subjectivity* (n=31). For the inherentness condition, participants rated the "essentiality" of adjective-noun object descriptions; they were told that they would be "asked to decide how essential the adjective is to the meaning of the noun which it modifies, that is, how substantive or inherent the adjective seems in its meaning." This language comes from Martin's original "substantiveness" task instructions (Martin, 1969, Expt. VII). Participants indicated their rating using a slider with endpoints labeled "completely nonessential" (coded as 0) and "completely essential" (coded as 1).

The subjectivity condition was a replication of our original "subjectivity" experiment, modified to match the inherentness task. Participants rated the "subjectivity" of adjective-noun object descriptions (and not adjectives in isolation; cf. our *Expt. 1: Subjectivity*); they were told that they would be "asked to decide how subjective the adjective seems in its meaning." Participants indicated their rating using a slider with endpoints labeled "completely objective" (coded as 0) and "completely subjective" (coded as 1).

Participants completed a total of 26 trials, one for each of the adjectives in Table 1; nouns were chosen at random from the list in Table 1. All 72 participants indicated that they were native speakers of English; we analyze their data below.

Results. We averaged the inherentness and subjectivity scores for each adjective. Fig. S.1 plots mean inherentness (*left*) and subjectivity (*right*) scores against the inferred distance naturalness scores from *Expt. 1: Ordering preferences*. As Fig. S.1 demonstrates, inherentness does little to predict ordering preferences, accounting for 0.0% of the variance in the naturalness ratings (r^2=0.00, 95% CI [0.00, 0.02]). However, we have replicated the success of adjective subjectivity, which accounts for 75% of the variation in the naturalness ratings (r^2=0.75, 95% CI [0.53, 0.84]).

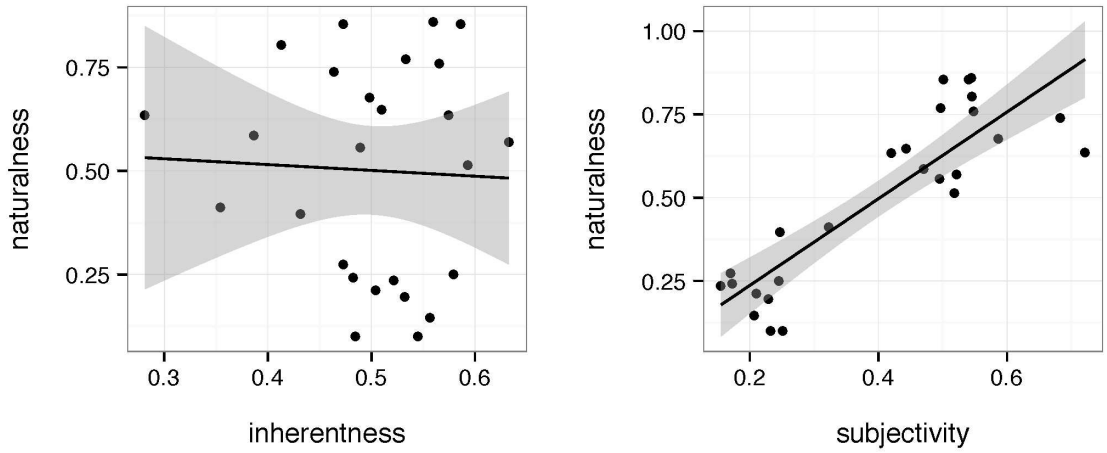

Figure S.1: Mean naturalness ratings plotted against mean inherentness (*left*) and subjectivity (*right*) scores for each of the 26 adjectives tested. Here and in all other graphs, gray ribbons represent 95% confidence intervals using a t-based approximation.

Discussion. Using a modified version of Martin's (1969) task measuring adjective inherentness, we failed to find *any* correlation between inherentness and ordering preferences for our set of 26 adjectives. Indeed, our findings match Martin's own, namely that inherentness is a poor predictor of adjective ordering preferences. We also measured adjective subjectivity, which allows for a direct comparison of the two predictors: whereas inherentness accounted for no variation, subjectivity accounted for 75% of the variation in our ordering preferences. Despite the many claims to its success in the literature (e.g., Sweet, 1898; Whorf, 1945; Kemmerer, 2000), adjective inherentness does little (or nothing) to explain the observed regularities in adjective order. However, subjectivity continues to predict ordering preferences.

2 Subjectivity vs. subsectivity

Next, we turn to a more abstract but no less popular factor meant to account for ordering preferences: the logical distinction between intersectivity and subsectivity. Semanticists argue that adjectives split on the basis of how they modify the nouns with which they compose (for discussion, see Kamp and Partee, 1995). Some adjectives are *intersective*: the outcome of their modification is the intersection of the nominal denotation with the adjectival denotation, as in the following example, where *carnivorous mammal* describes those things that hold the property of being carnivorous and of being a mammal.

$$\begin{aligned} [\![\text{carnivorous}]\!] &= \{x : \mathit{carnivorous}(x)\} \\ [\![\text{mammal}]\!] &= \{x : \mathit{mammal}(x)\} \\ [\![\text{carnivorous mammal}]\!] &= \{x : \mathit{carnivorous}(x) \ \&\ \mathit{mammal}(x)\} \\ &= [\![\text{carnivorous}]\!] \cap [\![\text{mammal}]\!] \end{aligned}$$

Other adjectives are *subsective*: rather than intersecting with a nominal denotation, the outcome of subsective modification is a subset of the nominal denotation, which is used to determine the comparison class for the adjective. Subsective adjectives depend on the nominal with which they compose to fix their meaning. Take *skillful*: a skillful surgeon is not necessarily a skillful violinist, but a skillful surgeon is necessarily a surgeon:

$$[\![\text{skillful surgeon}]\!] \subseteq [\![\text{surgeon}]\!]$$

Finally, there are those adjectives that are neither clearly intersective nor clearly subsective. The most obvious cases are so-called *privative* adjectives like *fake*, which require that the outcome of modification is a proper complement of the nominal denotation: a fake gun is not a gun. More generally, these non-intersective, non-subsective adjectives preclude the inference that the things described hold the property named by the noun:

$$[\![\text{fake gun}]\!] \not\subseteq [\![\text{gun}]\!]$$
$$[\![\text{former senator}]\!] \not\subseteq [\![\text{senator}]\!]$$

What matters for present purposes is that some authors assume that the means of modification determines adjective order, such that intersective adjectives occur closer to the noun than subsective adjectives. We find the clearest statement of this claim in Truswell (2009), who assumes that hierarchical dominance leads to linear precedence:

> 1. Subsective adjectives dominate intersective adjectives.
> 2. Modal [(i.e., non-subsective)] adjectives are freely ordered with respect to subsective and intersective adjectives, although they tend to dominate both classes.

According to Truswell, intersective adjectives occur closest to the modified noun, subsective adjectives occur farther than intersective adjectives, and non-subsective adjectives are freely ordered, despite their tendency to occur farthest of all.[2]

Subsectivity is not a true competitor to the subjectivity hypothesis, in two ways. Most importantly, the intersective–subsective distinction is a binary one, but there is systematic *graded* behavior in the ordering of adjectives. The binary distinction could at best explain the first bit of ordering preference data (though this could be a useful start). Second, the intersective–subsective distinction is a theoretical construct, not a natural human concept; it takes consultation with a trained linguist to fix the status of a given adjective.[3] This makes it theoretically useful, but behaviorally problematic. Still, we attempted to compare the predictions of the two proposals. To do so, we coded our adjectives as either *subsective*,

[2]See Sproat and Shih (1991) for a similar claim, which uses the language of the "relative" (i.e., subsective) vs. "absolute" (i.e., intersective) distinction (cf. Siegel, 1976).

[3]To see what it would take to operationalize the distinction as a behavioral measure, see our attempt at operationalizing concept formability in Section 3.2 below.

intersective, or *other* using Truswell's classifications: shape, color, nationality, and material adjectives were coded as *intersective*; all other adjectives were coded as *subsective*, except for the non-intersective class "X" adjectives from Expt. 2, which we coded as *other*.

The first thing to note is that the two predictors (i.e., subjectivity and subsectivity) are extremely highly correlated (Expt. 1: r^2=0.89, 95% CI [0.74, 0.94]; Expt. 2: r^2=0.52, 95% CI [0.35, 0.65]). This correlation should come as no surprise, given that the intersective–subsective distinction hangs on the context sensitivity of the adjective: intersective adjectives have fixed interpretations (i.e., they are maximally objective), while the interpretation of subsective adjectives varies with the modified noun. Given this high correlation and the observed success of subjectivity in predicting ordering preferences, we expect subsectivity to do well in predicting ordering preferences. However, the binary nature of the subsectivity predictor cannot account for the variation *within* the classes of subsective and intersective adjectives.

Fig. S.2 plots adjective subjectivity scores against naturalness ratings from Expt. 1 (*top*) and Expt. 2 (*bottom*), grouping adjectives by subsectivity class. To evaluate the role of subjectivity *over and above subsectivity* in predicting ordering preferences, we performed a nested linear model comparison. The models we compared predicted naturalness ratings by SUBSECTIVITY only, or by SUBSECTIVITY and SUBJECTIVITY. The model comparison revealed that subjectivity does explain additional variance in ordering preferences beyond the intersective–subsective distinction (Expt. 1: $F(1, 2337) = 25.28, p < 0.001$; Expt. 2: $F(1, 23786) = 361.78, p < 0.001$): as Fig. S.2 shows, even within the subsectivity classes, subjectivity continues to predict ordering preferences. Thus, we continue to find support for the subjectivity hypothesis. Performing the comparison in the reverse, we see that subsectivity also explains variance over and above subjectivity (Expt. 1: $F(1, 2337) = 36.89, p < 0.001$; Expt. 2: $F(2, 23786) = 242.60, p < 0.001$). That is, both subjectivity and subsectivity explain independent variance in ordering preferences—a finding which warrants future study.

3 Testing compositional accounts

Compositional accounts of ordering preferences hold that the fundamental factor in predicting adjective ordering is whether or not an adjective is used to form a complex concept/subkind description: first you form the concept, then you modify it with additional adjectives (McNally and Boleda, 2004; Svenonius, 2008).[4] This would imply that an interaction between the noun and a modifying adjective—whether they combine to form a complex concept—should have a large influence on adjective ordering. Indeed, a more general hypothesis is that *some* interaction between a noun and adjective will influence how closely the adjective is placed to that noun. This interaction could be caused by concept-formation, differential subjectivity, or other factors. We tested for such an interaction in Expt. 1 and in Expt. 2, and found that noun-specific naturalness did not explain any variance in ordering preference above and beyond adjective-level naturalness. Here we follow up on this result, first with a new set of materials that were chosen to maximize the probability of noun effects, and then with our attempt at operationalizing a concept-formability metric.

[4]Bouchard (2005) makes a similar claim, namely that the formation of complex concepts can override adjective ordering preferences.

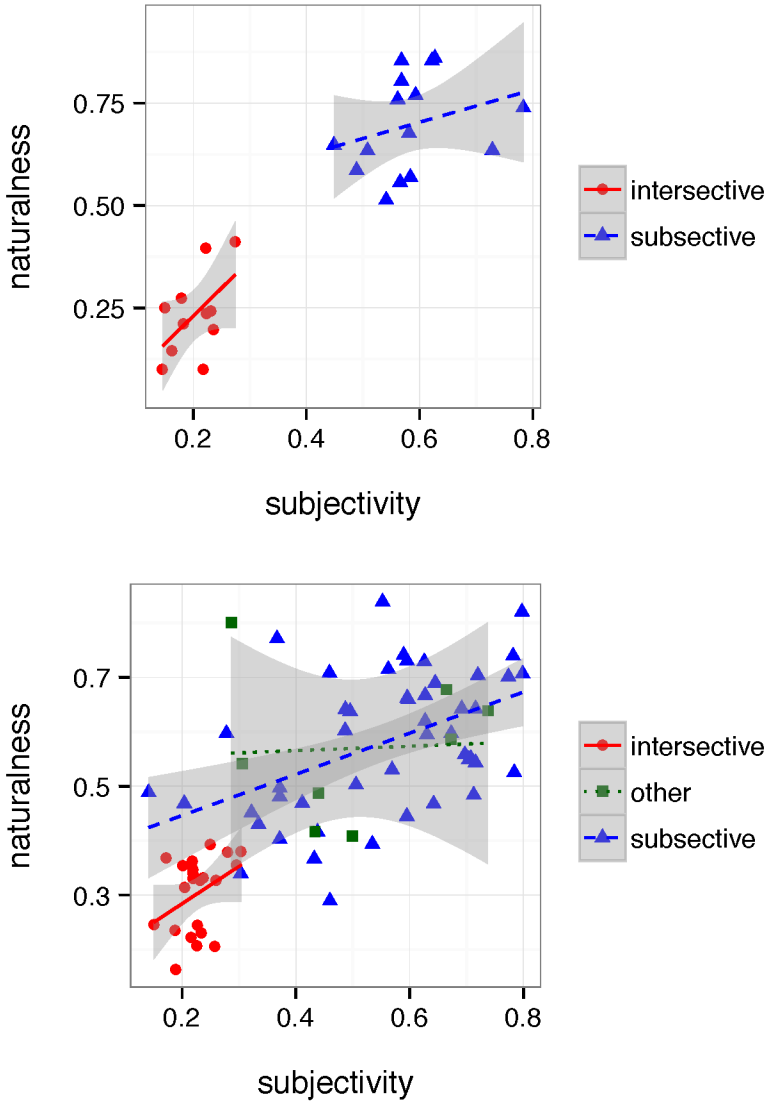

Figure S.2: Mean naturalness ratings plotted against mean subjectivity scores grouped by subsectivity for the original set of 26 adjectives tested in Expt. 1 (*top*) and the expanded set of 78 adjectives tested in Expt. 2 (*bottom*).

3.1 In search of noun effects

This experiment was a direct replication of *Expt. 1.1 Ordering preferences*, using a different set of nouns. We aimed to choose nouns that formed idiomatic, complex concepts with our set of 26 adjectives (along the lines of *bad apple* or *blue cheese*). Complex concepts tend to be described using the two-word name, yielding more occurrences of this bigram than would be expected from the unigram frequencies of the noun and adjective. This provided a way to extract candidate complex concepts from corpora.

Participants. We recruited 50 participants through Amazon.com's Mechanical Turk crowdsourcing service. Participants were compensated for their participation.

Design and methods. The design was identical to our original naturalness ratings experiments: participants were asked to indicate which of two object descriptions sounded more natural, using a sliding scale. Each description featured a noun modified by two adjectives; description pairs contained the same words with the relative adjective order reversed (e.g., "the big blue thing" vs. "the blue big thing"). Adjectives were chosen at random from the set in Table 1. The nouns were a smaller set of five (compared to the original ten). Nouns were chosen to maximize the probability of detecting noun-specific effects on adjective ordering preferences. In particular, we expected that nouns that are likely to form complex concepts should be highly collocational with that adjective. We thus searched for nouns that occur in particular adjective-noun phrases more frequently than predicted by the individual noun and adjective probabilities; in other words, nouns whose adjective-noun combinations were under-predicted by their individual word probabilities.

To find these nouns, we estimated the probability $p(A)$ of each adjective from our set of 26 by computing its relative frequency in an adjective-noun sequence in the BNC. We then computed the relative frequency of each noun $p(N)$ occurring in an adjective-noun sequence. Finally, we estimated the predicted joint probability of each adjective-noun combination by taking the product of each individual probability estimate: $\hat{p}(A, N) = p(A) \cdot p(N)$. Comparing $\hat{p}(A, N)$ to the empirically estimated $p(A, N)$ establishes which adjective-noun combinations are under-predicted—more collocational—and thus likely to name complex concepts. We then restricted nouns to those 50 that maximize the observed range of under-predictedness while simultaneously requiring that each noun be attested to occur with at least 11 of the 26 adjectives; from these 50 nouns, we selected the following four: *apple, cheese, eyes, hair*.[5] (Recall that *cheese* occurred in our original materials, where it suggested possible by-noun effects with the adjectives *hard* and *soft*.) To these four nouns we added a fifth: *thing*. While *thing* did not occur in the top 50, it did occur naturalistically with the most adjectives (23) out of the set of 26, thus allowing it to serve as a filler for the various object descriptions. The selected nouns, together with the number of adjectives they occur with, their range of ratios of empirical to predicted joint probabilities, and their minimum / maximum ratios, are shown in Table S.1.

[5] We restricted to a small subset of the 50 target nouns in order to maximize statistical power to identify noun effects on ordering.

Noun	# of adjectives	range of ratios	minimum ratio	maximum ratio
thing	23	10.4	0.1	10.5
eyes	18	120.6	0.12	120.7
hair	15	82.9	0.03	83.0
cheese	13	114.0	0.4	114.4
apple	11	674.0	1.1	675.1

Table S.1: For each chosen noun, the number of adjectives (out of 26) that it occurs with; and for each adjective A that the noun occurs with, the range of ratios $p(A, N)/\hat{p}(A, N)$ (empirical to predicted probability of occurrence); the minimum ratio; and the maximum ratio.

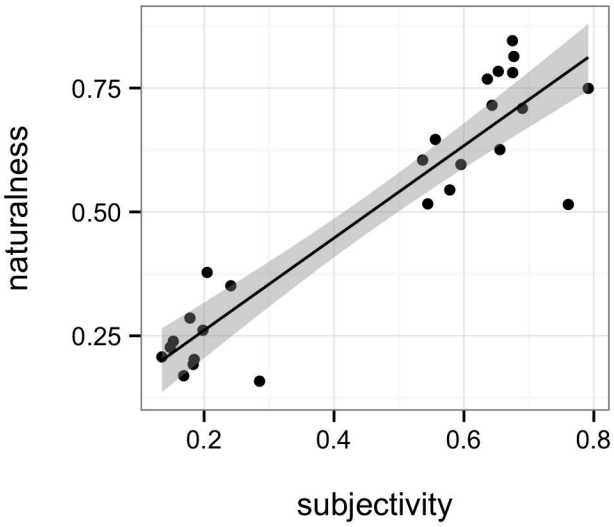

Figure S.3: Mean naturalness ratings plotted against mean subjectivity scores for each of the 26 adjectives tested in the new order preference experiment.

Results. To evaluate the role of specific noun information in determining ordering preferences, we performed the same nested linear model comparison from our original naturalness ratings experiment. The models we compared predicted naturalness ratings either by ADJECTIVE (i.e., the adjective farthest from the noun) only, or by ADJECTIVE together with its interaction with NOUN (i.e., the modified noun). The model comparison revealed that noun-specific ratings did not explain any additional variance in ordering preference beyond adjective-level ratings ($F(104, 2366) = 1.07, p < 0.30$). Thus, we again fail to find evidence of noun-specific effects on ordering preferences in our new materials.

Given the lack of noun effects on ordering preferences, we should continue to find that adjective subjectivity predicts ordering preferences. Indeed, it does: adjective subjectivity scores (obtained in *Expt 1.2 Subjectivity*) account for 85% of the variance in the new naturalness ratings (r^2=0.85, 95% CI [0.64, 0.93]; Fig. S.3). As with our original materials, more subjective adjectives are preferred farther from the noun.

Discussion. Using nouns chosen to maximize the probability of complex concepts formed with our 26 adjectives, we failed to find evidence that adjective ordering preferences depend on the modified noun. We do continue to find that subjectivity predicts ordering preferences. Thus, we failed to find evidence in support of compositional accounts of ordering preferences, which hold that the most important factor in determining order is whether or not an adjective forms a complex concept with the noun it modifies. It remains possible that noun-effects (and hence effects of concept composition) would be found with a different set of adjectives. Moreover, there could be noun effects that are not related to concept composition. However, this already suggests that the effect of modified nouns explains only a small part of the overall story of adjective order, while subjectivity is a robust predictor.

3.2 Operationalizing concept-formability

Failing to find support for the most general prediction of the compositional account, namely an effect of nouns on ordering preferences, we tried a more targeted approach: operationalizing a concept-formability metric, and testing its predictions on the adjective ordering preferences that we measured in *Expt. 1: Ordering preferences*. As with the studies in our paper, the work lies in operationalizing the abstract notion of whether or not an adjective-noun combination tends to form a complex concept. The literature on the topic presupposes that intuitions about concept formability are systematic and generalizable (McNally and Boleda, 2004; Svenonius, 2008); the closest we found in these papers to a proposal for an empirical measure of this factor is the following distinction.

According to McNally and Boleda, the key issue is one of entailment. When an adjective modifies a noun intersectively, the objects described hold both the property named by the noun and the property named by the adjective: a "male architect" is both male and an architect (McNally and Boleda, 2004, p. 179, ex. 2). When an adjective and a noun combine to form a complex concept (i.e., a subkind description), the objects described hold the property named by the noun, but not necessarily the property named by the adjective; the modification is (ostensibly) subsective. The authors give the Catalan example *arquitecte tècnic* "technical architect," which names a kind of architect but not necessarily technical things (McNally and Boleda, 2004, p. 179, ex. 1; cf. the discussion of *wild rice* in Svenonius, 2008).[6] Using our original set of materials, we tested whether the objects named by an adjective-noun description hold 1) the property named by the adjective, and 2) the property named by the noun.

Participants. We recruited 40 participants through Amazon.com's Mechanical Turk. Participants were compensated for their participation.

Design and methods. Participants were presented with 26 adjective-noun object descriptions. They were instructed to "consider the things that might be described" as `adjective-noun` (e.g., "round desks"), then rate how likely it is that the things described 1)

[6]This distinction should ring familiar from the discussion of intersective vs. subsective modification above. Indeed, as McNally and Boleda characterize it, concept formability is closely related to the intersective/subsective distinction, and the current experiment operationalizes both.

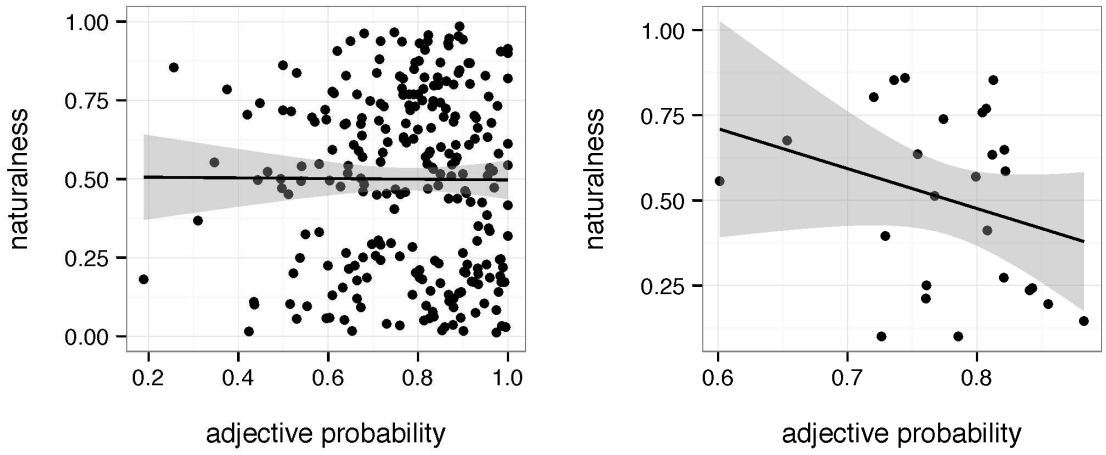

Figure S.4: Mean naturalness ratings plotted against mean adjective probability (i.e., concept-formability) scores for each of the 26 adjectives tested. In computing mean adjective probability, we either took into account (*left*) or collapsed over (*right*) specific noun information.

hold the `adjective` property (e.g., "Are those things round?"), and 2) hold the `noun` property (e.g., "Are those things desks?"). Participants indicated their probability ratings using sliders with endpoints labeled with "definitely not" (coded as 0) and "definitely" (coded as 1). All 40 participants indicated that their native language was English; we report their results below.

Results. To ensure that we did not wash out the noun-specific effects that this analysis was supposed to pick up, we computed two sets of adjective probability scores: 1) adjective-noun probability scores that simply averaged ratings across participants, and 2) adjective probability scores that collapsed over the modified noun. Fig. S.4 plots the adjective-noun (*left*) and adjective (*right*) scores against the naturalness ratings from *Expt. 1: Ordering preferences*. According to the predictions of the compositional account, lower adjective (or adjective-noun) probability should lead to lower naturalness (i.e., a decreased preference to place the adjective farther from the noun): as the property named by an adjective is less likely to apply straightforwardly to the objects named, the probability that the adjective forms a complex concept with the noun it modifies increases. While we do observe this trend for the simple adjective scores in Fig. S.4 (*right*), we also see that adjective probability scores predict just 8% of the variance in our preference data ($r^2 = 0.08$; 95% CI [0.00, 0.33]). The adjective-noun scores perform even worse, accounting for none of the variance in the adjective-noun preference data ($r^2 = 0.00$; 95% CI [0.00, 0.00])

Discussion. Recall that at its worst, subjectivity predicts 70% of the variance in our preference data. While it is quite possible that concept-formability plays an important role for some cases (such as *arquitecte tècnic*), we did not find evidence that it was critical to ordering preferences in our broad set of items.

4 Conclusion

We have compared the subjectivity hypothesis with the predictions of three alternative accounts of adjective order: inherentness, subsectivity, and concept-formability. Our results demonstrate the robust replicability of the success of subjectivity in predicting adjective order. The results also demonstrate that the alternative accounts we considered add little to our understanding of ordering preferences: everywhere we looked, subjectivity continued to best predict ordering preferences. It is certainly possible that our operationalizations have missed key aspects of the theoretical proposals, or that the predicted relationships could be found with different materials. But at the least, this exploration has led us to better appreciate the intuitive appeal of the subjectivity hypothesis, which targets an easily-accessed, stable psychological construct: adjective subjectivity.

References

Bouchard, D. (2005). Sériation des adjectifs dans le SN et formation de concepts. *Recherches linguistiques de Vicennes 34*, 125–142.

Kamp, H. and B. Partee (1995). Prototype theory and compositionality. *Cognition 57*, 129–191.

Kemmerer, D. (2000). Selective impairment of knowledge underlying prenominal adjective order: evidence for the autonomy of grammatical semantics. *Journal of Neurolinguistics 13*, 57–82.

Martin, J. (1969). Semantic determinants of preferred adjective order. *Journal of Verbal Learning and Verbal Behavior 8*, 697–704.

McNally, L. and G. Boleda (2004). Relational adjectives as properties of kinds. *Empirical Issues in Formal Syntax and Semantics 5*, 179–196.

Siegel, M. E. A. (1976). *Capturing the Adjective*. Ph. D. thesis, University of Massachusetts Amherst.

Sproat, R. and C. Shih (1991). The cross-linguistic distribution of adjective ordering restrictions. In C. Georgopoulos and R. Ishihara (Eds.), *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, pp. 565–593. Dordrecht: Kluwer Academic Publishers.

Svenonius, P. (2008). The position of adjectives and other phrasal modifiers in the decomposition of DP. In L. McNally and C. Kennedy (Eds.), *Adjectives and Adverbs: Syntax, Semantics, and Discourse*, pp. 16–42. Oxford: Oxford University Press.

Sweet, H. (1898). *A New English Grammar, Logical and Historical*. London: Clarendon Press.

Truswell, R. (2009). Attributive adjectives and nominal templates. *Linguistic Inquiry 40*, 525–533.

Whorf, B. L. (1945). Grammatical Categories. *Language 21*(1), 1–11.

Wulff, S. (2003). A multifactorial corpus analysis of adjective order in English. *International Journal of Corpus Linguistics 8*(2), 245–282.