

Adjective Ordering

Gregory Scontras, Judith Degen, and Noah D. Goodman

June 24, 2015

Abstract

Here is the abstract

Contents

1	Introduction	1
1.1	English	2
1.2	Cross-linguistic stability	2
2	Corpus work	2
3	Experiments	3
3.1	Experiment 1: Subjectivity	3
3.1.1	Participants	3
3.1.2	Design and methods	3
3.1.3	Predictions	3
3.1.4	Results	3
3.1.5	Discussion	3
3.2	Experiment 2: Ordering preferences	3
3.2.1	Participants	3
3.2.2	Design and methods	4
3.2.3	Predictions	4
3.2.4	Results	4
3.2.5	Discussion	6
4	Possible accounts	6
4.1	Memory	6
4.2	Syntax	6
4.3	UID	7
4.4	Flexibility	7
5	General discussion	7

1 Introduction

Background on adjective ordering preferences

1.1 English

1.2 Cross-linguistic stability

2 Corpus work

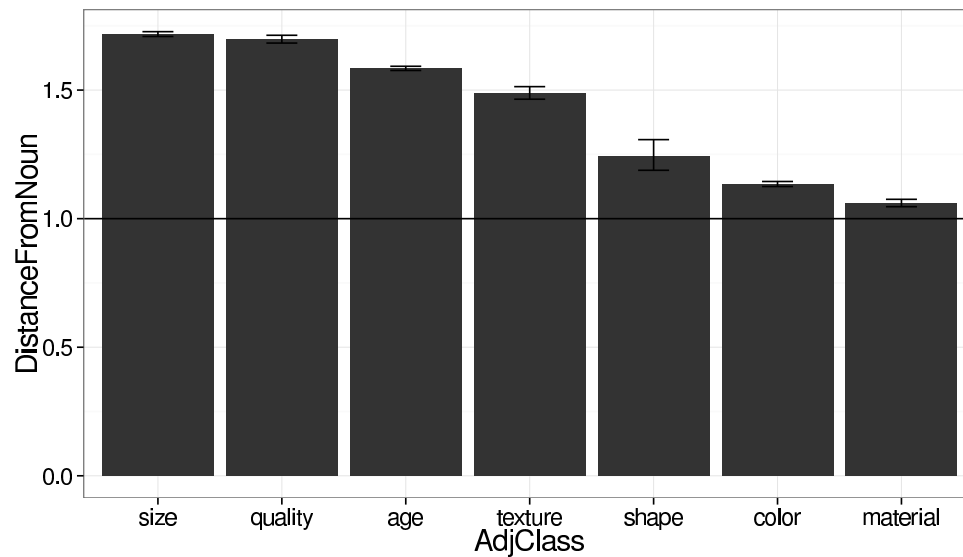


Figure 1: Average distance from noun by adjective class for cases with at least two modifying adjectives (39,199 cases).

Based on the average distance-from-noun scores calculated in Fig. 1, we may infer the ordering preferences in (1).

- (1) *Adjective order preferences inferred from corpus counts (Fig. 1):*
 size \geq quality > age > texture > shape > color > material

Table 1: Adjectives and nouns used in experimental stimuli.

ADJECTIVE	CLASS	ADJECTIVE	CLASS	NOUN	CLASS
good	quality	old	age	apple	food
bad	quality	new	age	banana	food
round	shape	rotten	age	carrot	food
square	shape	fresh	age	cheese	food
big	size	red	color	tomato	food
small	size	yellow	color	chair	furniture
huge	size	green	color	couch	furniture
tiny	size	blue	color	fan	furniture
short	size	purple	color	TV	furniture
long	size	brown	color	desk	furniture
smooth	texture	wooden	material		
hard	texture	plastic	material		
soft	texture	metal	material		

3 Experiments

3.1 Experiment 1: Subjectivity

3.1.1 Participants

3.1.2 Design and methods

3.1.3 Predictions

3.1.4 Results

3.1.5 Discussion

3.2 Experiment 2: Ordering preferences

We have so far established a connection between adjective subjectivity (measured by faultless disagreement scores in Expt. 1) and adjective proximity to nouns (obtained from corpus counts in Section 2). Our next task is to verify the adjective ordering preferences that we have inferred from the corpus and from the literature on the topic, and to evaluate the predictive power of subjectivity in determining adjective order. To that end, we elicited naturalness judgments on adjective-adjective-noun object descriptions, permuting the relative order of adjectives. We then compared these naturalness judgments with the faultless disagreement scores from the previous experiment.

3.2.1 Participants

We recruited 50 participants through Amazon.com’s Mechanical Turk crowd-sourcing service. Participants were compensated for their participation.

3.2.2 Design and methods

Participants were asked to indicate which of two descriptions of an object sounded more natural. Each description featured a noun modified by two adjectives, for example “the red small chair” or “the small red chair”. Description pairs contained the same words, with relative adjective order reversed. Descriptions were random combinations of two adjectives and a noun from the list in Table 1 (compiled via the procedure described in Section 2), with the constraint that no description contained adjectives from the same adjective class. On each trial, participants indicated which description sounded more natural by adjusting a slider whose endpoints were labeled with the competing descriptions; an example trial appears in Fig. 2.

Figure 2: Example trial from Expt. 1; participants indicated the more natural of two adjective-adjective-noun descriptions on a sliding scale.

Only native speakers of English with IP addresses located within the United States were included in the analyses; we analyzed data from 45 participants.

3.2.3 Predictions

3.2.4 Results

For each pair of adjective classes, we determine the preferred ordering on the basis of the ratio of ratings for the each order. Ratio scores greater than 1 indicate the preferred ordering; these ratio scores for the preferred orderings are plotted in Fig. 3.

On the basis of these preferred orderings, we may calculate the average distance from noun for each class, plotted in Fig. 4. Comparing Fig. 4 with the corpus-based distance scores reported in Fig. 1 confirms the reliability of the current paradigm: we replicate near exactly the qualitative order of adjective class distance from noun.

We may infer the order preferences in (2) (cf. the preferences inferred from corpus counts in (1) above).

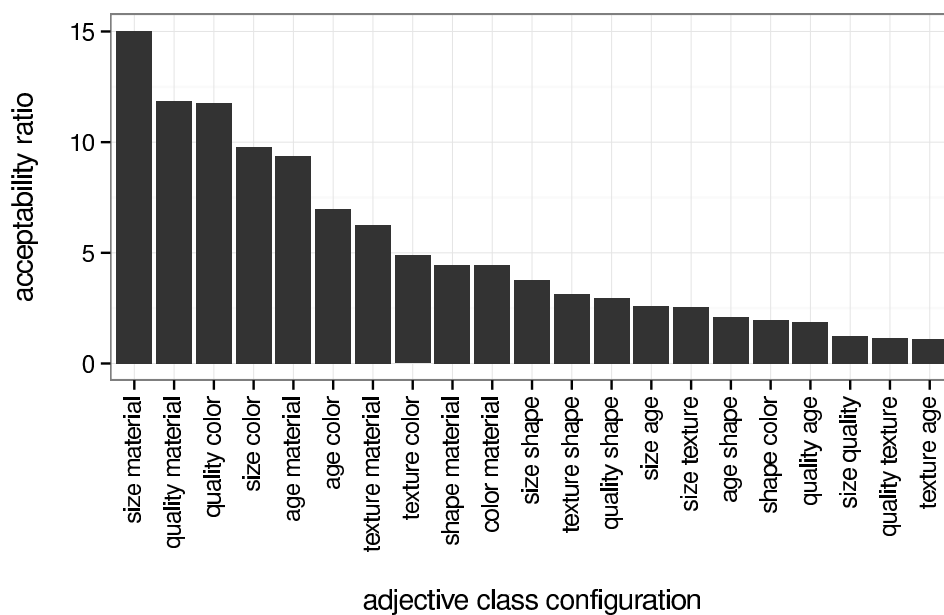


Figure 3: Acceptability ratios for the preferred adjective class orderings.

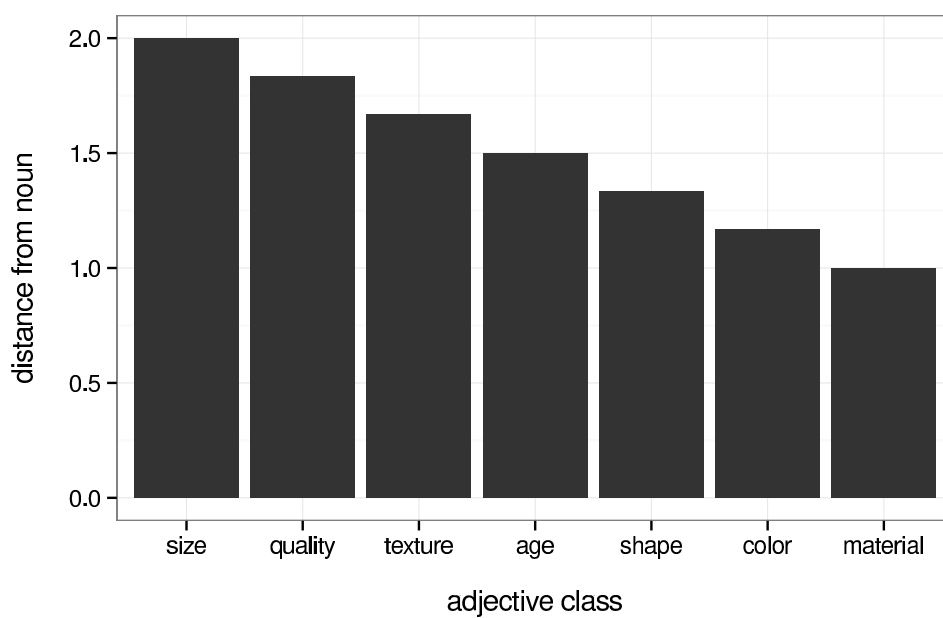


Figure 4: Average distance from noun for each adjective class calculated from order preference ratings.

- (2) *Adjective order preferences inferred from acceptability ratings (Fig. 4):*
 size > quality > age > texture > shape > color > material

In addition to calculating preference ratios by class configurations, we also calculated preferred

orderings for all of the specific adjective pairings. On the basis of how often an adjective from a given class occurred first in an adjective-adjective-noun configuration, we may infer the relative distance from the noun a given class prefers. Fig. 5 plots these average distance scores, where a value of 2 signals that a class’s adjectives always occur first in preferred adjective-adjective-noun orderings, and a value of 1 indicates that a class’s adjectives always occur second.

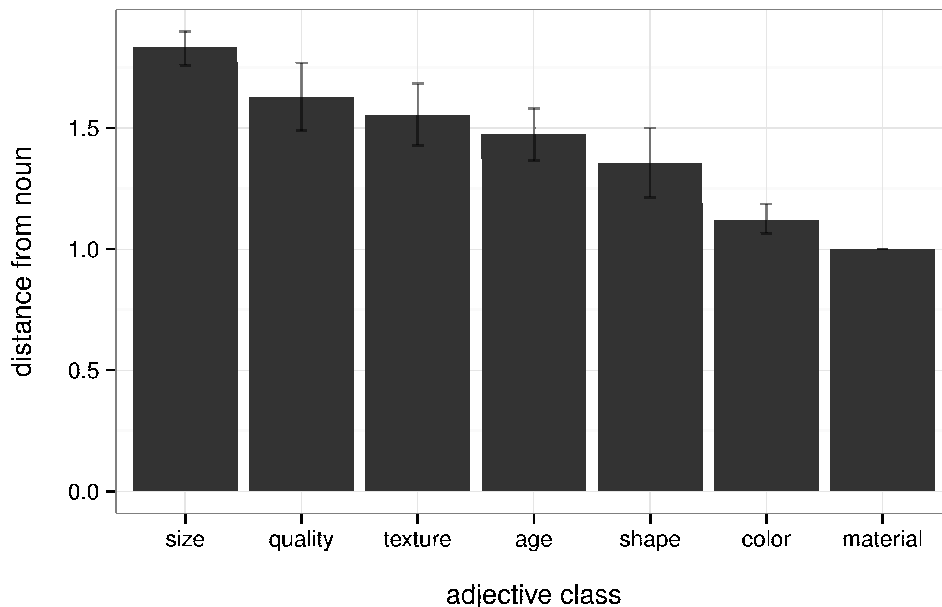


Figure 5: Average distance from noun for each adjective class by computing how often adjectives from that class occur first in preferred adjective-adjective-noun orderings.

Finally, and most directly related to our hypothesis concerning adjective subjectivity in ordering preferences, we compared acceptability ratings from the current experiment with faultless disagreement scores from Expt. 1. To do so, we first calculated a difference score for each class configuration, $\text{CLASS1} - \text{CLASS2}$, subtracting the average faultless disagreement score for CLASS1 from the average faultless disagreement score for CLASS2. Fig. 6 plots class configuration acceptability ratings against faultless disagreement difference scores.

3.2.5 Discussion

4 Possible accounts

4.1 Memory

Forgetting account

4.2 Syntax

John Beavers spray/load work

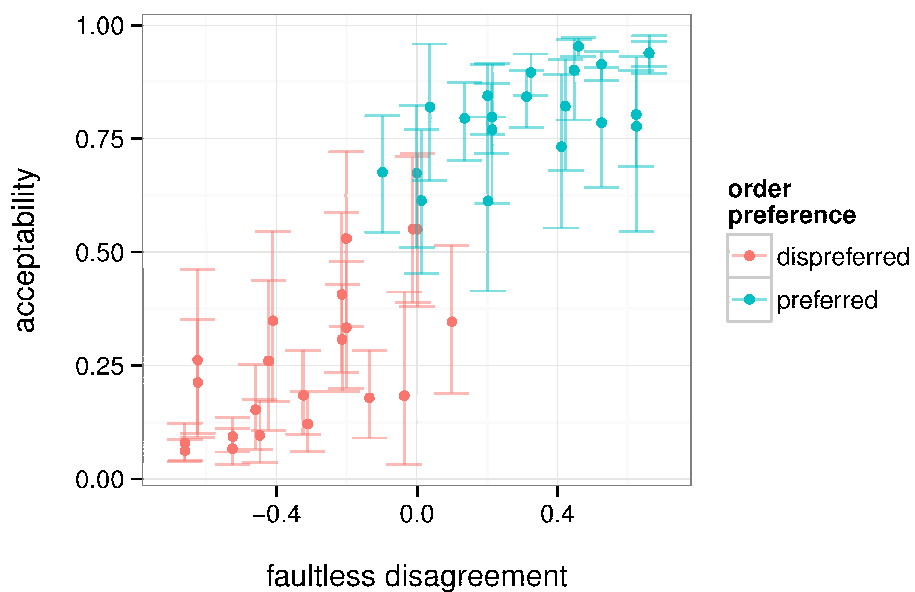


Figure 6: By-class adjective order preferences plotted against difference in faultless disagreement.

4.3 UID

Different notions of informativity

4.4 Flexibility

Blue is less subjective so I rarely speak of “blue for X”, whereas more subjective and flexible *good* is often spoken of in terms of “good for Y”.

5 General discussion