

On the grammatical source of adjective ordering preferences

Gregory Scontras

University of California, Irvine

Abstract

Scontras et al. (2017) present experimental evidence demonstrating that the best predictor of adjective ordering preferences in the English noun phrase is the subjectivity of the property named by any given adjective: less subjective adjectives are preferred linearly closer to the nouns they modify. The current work builds on this empirical finding by proposing that the reason subjectivity predicts adjective ordering preferences has to do with the hierarchical structure of nominal modification. Adjectives that are linearly closer to the modified noun are often structurally closer, composing with the noun before adjectives that are farther away. Pressures from successful reference resolution dictate that less subjective, more useful adjectives contribute their meaning to the resulting nominal earlier, in an attempt to more effectively limit the reference search space.

Keywords: adjective ordering, subjectivity, hierarchical structure, modification, reference resolution

1. Introduction

Adjective ordering preferences determine the relative order of adjectives in multi-adjective strings. Such preferences dictate that *the small brown cardboard box* sounds much more natural than *the brown cardboard small box*, or any other ordering of the adjectives. These preferences are robustly attested, not only in English, but in a host of additional, unrelated languages (Dixon, 1982; Hetzron, 1978; LaPolla and Huang, 2004; Martin, 1969b; Sproat and Shih, 1991). Remarkably, the *same* preferences surface in each case. Even more remarkably, in post-nominal languages where adjectives follow the nouns they modify, the preferences are the mirror image of what are found in pre-nominal languages like English (Dixon, 1982; Hetzron, 1978; Sproat and Shih, 1991); at issue is the relative distance of an adjective from the noun it modifies.

Given their stability within and across languages, a glaring question presents itself: what factors determine these robust preferences? Answers to this question

*E-mail address: g.scontras@uci.edu, URL: <http://linguistics.uci.edu/scontras/>

stand to inform not only the preferences themselves, but also the psychological and grammatical systems from which these preferences emerge. For this reason, adjective ordering preferences have been the subject of targeted inquiry since Sweet (1898) wrote about them over a century ago. Hypotheses abound, ranging from the psychological (Sweet, 1898; Whorf, 1945; Ziff, 1960; Martin, 1969a) to the grammatical (Cinque, 1994; Scott, 2002; Laenzlinger, 2005; McNally and Boleda, 2004; Svenonius, 2008; Truswell, 2009). Still, significant progress has proven elusive, owing to the complex, careful empirical work required to test these hypotheses.

Recently, Scontras et al. (2017) brought behavioral and corpus data to bear on the question of adjective ordering. Distilling the proposals that preceded them, Scontras et al. advanced the hypotheses that property subjectivity determines the relative order of adjectives in multi-adjective strings, such that less subjective adjectives occur linearly closer to the nouns they modify (see also Quirk et al., 1985; Hetzron, 1978; Dixon, 1982; Tucker, 1998; Hill, 2012). In *the small brown cardboard box*, *cardboard* is much less subjective than *brown* or *small*, so *cardboard* is preferred closer to the modified noun.

With strong empirical footing for the factor determining ordering preferences, the question now shifts to precisely *why* subjectivity should play such a central role in adjectival modification. The remainder of this short paper proceeds as follows. Section 2 reviews the empirical methodology and findings from Scontras et al. (2017). Section 3 explores potential explanations of the empirical findings. Section 4 offers a proposal tying subjectivity to reference resolution and the hierarchical structure of adjectival modification. Section 5 concludes.

2. Subjectivity predicts adjective ordering preferences

To identify the factors at play in adjective ordering preferences, we must first determine what the preferences are that need explaining. To that end, Scontras et al. (2017) began their study by establishing a behavioral measure of ordering preferences. They asked experimental participants to indicate the preferred ordering of adjectives in adjective-adjective-noun strings (e.g., *the small brown chair* vs. *the brown small chair*). From these relative preferences the authors then calculated a single preferred-distance measure for each adjective tested, corresponding to that adjective’s overall proximity to modified nouns. The authors evaluated their behavioral proximity measure against naturalist productions from English corpora. Finding an extremely strong correlation between the behavioral measure and the corpus counts, the authors concluded 1) that naive speakers have reliable and robust adjective ordering preferences, and 2) that the behavioral measure faithfully captured these preferences. With a clear handle on the preferences, the authors then shifted their focus to the aspect of adjective meaning that they hypothesized best predicted ordering preferences: subjectivity.

Inspired by the various proposals about aspects of adjective meaning explaining their relative order in multi-adjective strings, Scontras et al. distilled past proposals into the intuitive psychological construct of subjectivity. Crucially, the

authors *operationalized* their subjectivity hypothesis as a behavioral measure for the purpose of empirical testing. Adjective subjectivity was measured by asking participants how “subjective” a given adjective was. These raw “subjectivity” scores were evaluated against a potentially more ecologically valid method: faultless disagreement (Kölbel, 2004; Barker, 2013; Kennedy, 2013; MacFarlane, 2014). To the extent that two speakers can disagree about a given property for an object without one speaker necessarily being wrong (e.g., disagreeing about whether or not a box counts as small), the property admits that degree of faultless disagreement, which stands proxy for the adjective’s subjectivity. Finding extremely high correlations between the “subjectivity” scores and the faultless disagreement measure, Scontras et al. concluded that both measures successfully capture adjective subjectivity.

With clear estimates of the preferences and of adjective subjectivity, the authors then tested the predictive power of subjectivity in explaining ordering preferences. Recall the hypothesis: adjectives with lower subjectivity scores are preferred linearly closer to the nouns they modify. Indeed, this is precisely what Scontras et al. found: an adjective’s semantics predicts its distance from the nouns it modifies, with subjectivity scores accounting for nearly all of the variance in the ordering preference data. Moreover, preference strength increased with the subjectivity differential. To get a clearer picture of the *relative* success of their subjectivity hypothesis, the authors then compared the predictions of subjectivity against operationalizations of competing proposals: adjective inherebntness (whereby adjectives with more “essential” meanings occur closer to modified nouns; Sweet, 1898; Whorf, 1945; Ziff, 1960), intersective vs. subsective modification (whereby intersective modifiers compose first in the hierarchical structure of nominals; Truswell, 2009), and concept formability (whereby adjectives that form complex, idiomatic concepts compose early; Bouchard, 2005; McNally and Boleda, 2004; Svenonius, 2008). In each case, subjectivity continued to be a better predictor of ordering preferences.

3. Why subjectivity?

Finding subjectivity to be a reliable and robust predictor, our task now is to explain why subjectivity should determine adjective ordering preferences: why should less subjective adjectives be preferred linearly closer to the nouns they modify? Scontras et al. hint at an answer in the discussion of their results, namely pressure from successful reference resolution. Before reviewing their discussion, it will be useful to first consider the range of possible answers to this *why* question. In the process, we also establish desiderata for successful answers.

To begin, we might propose that the observed subjectivity gradient emerges from a rigid syntax of adjectival modification: adjectives inhabit specialized syntactic projections depending on their semantic class (e.g., Color Phrase for color adjectives, Shape Phrase for shape adjectives, etc.; Cinque, 1994; Scott, 2002; Laenzlinger, 2005), and these projections happen to order in a way that tracks subjectivity. While a cartographic approach along these lines might help to

explain the observed behavior, it leaves unanswered the question of *why* subjectivity should matter in the ordering of adjectival projections. Also problematic is the rigidity introduced by a syntax that allows only one ordering for any string of adjectives. This rigidity predicts categorical ordering preferences, yet Scontras et al. observed graded judgments that track differential subjectivity. Thus, a cartographic syntax appears to be a nonstarter for explaining why subjectivity should predict ordering preferences.

Shifting our sights to psychological explanations, we might try to account for the role of subjectivity by making appeal to the relative salience of properties for the nouns they modify. Properties that are more salient—more *inherent* to the objects described by the noun (Sweet, 1898; Whorf, 1945; Biber et al., 1999)—ought to be more accessible in the construction of nominals. An account based on the accessibility of adjectives during the on-line construction of nominal phrases stands to extend straightforwardly to languages with post-nominal adjectives where the preferences are preserved in the reverse: phrases are built outward from their heads, so more accessible adjectives occur linearly closer to the head of the nominal construction: the noun. The leap that must be made, however, links subjectivity to inherentness and thereby to accessibility. Unfortunately, this leap appears untenable. Scontras et al. measured adjective inherentness and compared its predictions with those of subjectivity. Whereas subjectivity accounted for at least 75% of the variance in the ordering preferences, inherentness accounted for 0%. While an explanation in terms of adjective accessibility might still prove promising, the implementation via inherentness lacks empirical support.

Rather than inherentness, we might try tying adjective accessibility to adjective frequency, such that more frequent adjectives are more accessible during the hierarchical construction of nominal phrases. To account for the role of subjectivity, we would expect more frequent—and thus more accessible—adjectives to have lower subjectivity scores. Scontras et al. investigated the role of adjective frequency in ordering preferences, finding it to be a significant predictor. However, frequency applies pressure in the direction opposite to what we have been considering: in English, more frequent adjectives occur *farther* from the modified noun because they occur *linearly* early in multi-adjective strings (cf. Bock, 1982; Wulff, 2003). Moreover, the authors found that subjectivity continues to explain significant variance in the preferences over and above adjective frequency. While frequency likely contributes to the relative accessibility of adjectives during the construction of nominal phrases, its contribution is separate from that of subjectivity; the two forces work in tandem and in opposite directions.

We find perhaps the most thought-through version of the psychological accessibility hypothesis in Martin’s (1969b) experimental investigations. Inspired by prior results demonstrating the predictive power of “definiteness of denotation” in adjective ordering preferences (Martin, 1969a), Martin set out to test the hypothesis that adjectives occurring linearly closer to nouns indeed are more accessible. Participants completed a series of elicited production tasks in which they were shown visual arrays of objects and asked to name specific properties

of the objects they saw (e.g., size vs. color). There were two versions of the experiments: one for English speakers, and another for speakers of Indonesian, a post-nominal language with the mirror image of the English ordering preferences. By measuring production latencies, Martin discovered that adjectives preferred linearly closer to nouns are produced more quickly. From this he concluded that adjectives closer to the noun are more accessible than adjectives preferred farther away.

Before accepting Martin's results as unambiguous support for an adjective accessibility hypothesis, we must confront two issues. First, production latencies in context likely depend on more than the relative accessibility of words from memory; how can we be sure that the observed differences in latencies did not derive from low-level properties of the visual displays? If the issue at play is truly the *lexical* accessibility of adjectives, then the displays should have controlled for the relative perceptual salience of the properties being named. As things stand, we have no way of teasing apart lexical accessibility from visual salience in Martin's results. Second, if accessibility truly determines adjective ordering preferences, why should adjective frequency apply pressure in the opposite direction? Relative frequency surely determines adjective accessibility, and we saw above that adjective frequency is a significant predictor of ordering preferences. However, more frequent adjectives are preferred early in the linear structure of nominal phrases, not closer to the nominal head (at least not in pre-nominal languages like English). Accessibility as measured by adjective frequency seems to be delivering the wrong predictions.

There of course remains the possibility that adjective accessibility is not the primary link between subjectivity and ordering preferences. In an attempt to tie properties of language structure to general principles of cognition, Bever (1970) advanced the hypothesis that structural norms—among them, adjective ordering preferences—emerge from the human perceptual system. To see how perception could determine the preferred ordering of a string of adjectives, one must appreciate the task of the parser, at least as envisaged by Bever. Upon encountering the speech stream, the parser relies on various heuristics to efficiently identify constituents and associate them with an appropriate argument structure. In service of this task, the parser needs an identification mechanism for noun phrases: where do they begin, and where do they end?

According to Bever, the more “nounlike” the adjective, the closer it appears to the modified noun (see also Biber et al., 1999; Posner, 1986). Returning to the small brown cardboard box, *cardboard* is the most felicitous when used as a noun, *brown* slightly less so, and *small* the least of all (cf. Bever's examples (68) and (69)). Now, why should ordering adjectives according to their nounlike character ease the burden of the parser in its search for noun phrase boundaries? Bever proposes that a linear parser identifies the beginning of a noun phrase with the presence of a determiner. That same parser identifies the right edge of the noun phrase with the transition from a clearly nounlike element to a lexical item that is “less uniquely a noun” (Bever, 1970, p. 323). In other words, the primary cue to the right edge of a noun phrase is a salient decrease in nounlike character. If adjectives were randomly ordered with respect to their nounlike

character, the parser might mistakenly identify noun phrase boundaries, as in [*the cardboard*] [*brown box*]. These early errors identifying phrase boundaries would cascade into a total failure for the sentence parse.

Bever’s proposal offers an intuitive explanation of the pressures that determine adjective ordering, and it readily extends to handle the mirror-image preferences in post-nominal languages. The proposal even offers a promising connection to subjectivity: perhaps less subjective adjectives yield more-well-defined categories, which are more amenable to naming with nouns. Thus, subjectivity determines nounlike character, and nounlike character determines relative order in multi-adjective strings. Unfortunately, Bever’s proposal suffers a serious flaw: it lacks empirical support. In her corpus analysis of English ordering preferences, Wulff (2003) calculated nounlike character for a host of adjectives and demonstrated that it does little by way of predicting adjective order. What effect nounlike character does have on the linear order of adjectives applies pressure in the direction opposite of Bever hypothesis: more nounlike adjectives are marginally more likely to occur *farther* from the modified noun (Wulff, 2003, p. 255).

It would appear that we have arrived at an impasse: accessibility-based accounts struggle to explain the full range of data from both pre- and post-nominal languages; they also face a serious obstacle in the form of lexical frequencies. Bever’s perception-based account seems well-suited for pre- and post-nominal languages, but Wulff’s facts suggest the proposal is misguided. And all of these proposals lack a clear connection to adjective subjectivity. There remains another strategy, however, which shifts the explanation from how language users use adjectives to what adjectives do for language users. As we shall see, a functional account along these lines stands the best chance of explaining the role of subjectivity in determining ordering preferences.

For Seiler (1978), the clue to understanding adjective ordering preferences lies in the task that adjectives perform: determination. Noun phrases are inherently referential, whether to real-world objects or to well-defined concepts. In either case, determiners—demonstratives, articles, numerals, quantifiers, adjectives, prepositional attributes, relative clauses—contribute to nominal meaning in service of pinning down a referent. With this function in mind, Seiler identifies regularities in the linear order of determiners. First, “the range of head nouns for which a determiner D is potentially applicable increases with the positional distance of that determiner from the head noun N” (Seiler, 1978, p. 308). For the purposes of adjective ordering, the more nouns an adjective can felicitously describe, the farther that adjective will appear from the noun. Seiler explicitly links determiner applicability with property inherentness: less broadly-applicable, more special-purpose adjectives name properties that are more inherent to the modified noun. He gives the example of *rote hölzerne Kugeln* ‘red wooden balls’:

The semantic structure of *Kugeln* qua solid objects naturally implies material constitution of some sort; it implies—with a lesser degree of naturalness—some property in the color spectrum. To this

gradient decrease in natural semantic implication corresponds the normal word order in which the ‘determiner’ with the strongly implied property is closer to the head noun than the ‘determiner’ with the less strongly implied property. (Seiler, 1978, p. 309)

Properties implied by the nominal have meanings that are at least partially contained already within the nominal meaning, hence the implication (cf. the notion of mutual informativity; Futrell, 2017). Determiners that are less implied by the nominal will be more informative (i.e., unexpected) when encountered, with greater informativity leading to a greater potential of pinning down the intended referent. Thus, according to Seiler, “the potential of a determiner D for singling out the object referred to by the head noun N increases proportionally with the positional distance of D from N” (Seiler, 1978, p. 309). The ordering that results presumably follows from the desire to introduce the more informative, more useful elements early in the construction of a nominal.

Seiler’s first claim—that adjectives describing a broader set of nouns appear farther from the modified noun—finds empirical support in Wulff’s corpus analysis (Wulff, 2003, pp. 266–267). However, the implication of Seiler’s second claim—that speakers introduce more useful elements (for the purpose of determination) earlier—fails in the case of post-nominal languages with mirror-image preferences. This reasoning would hold that speakers in post-nominal languages save the most useful adjectives for last, which stands in direct conflict with the explanation for pre-nominal languages. Still, we should not abandon the functional account of adjective ordering altogether. In the following section, we consider a different proposal that preserves Seiler’s intuition that “in order to fully understand the regularities we must look behind the mere facts and try to see the program and ultimately the purposive functions of which they are manifestations” (Seiler, 1978, p. 325).

4. Linking subjectivity to the hierarchical structure of modification

Let us begin as Seiler did, with the observation that adjectives aid in establishing reference. Starting with a noun like *box*, potential referents include every box in the discourse context. Where there are multiple boxes, the listener’s task of establishing reference amounts to a game of chance. We increase our odds of winning this game as we narrow down, or *determine* the set of potential referents. Encountering an adjective like *cardboard*, we now only consider the subset of boxes that are cardboard. Encountering *brown*, we limit ourselves to the cardboard boxes that are brown. Encountering *small*, we further limit ourselves to just those brown cardboard boxes that are small. From the set of all boxes we home in on the small brown cardboard boxes, a much smaller set indeed.

When it comes to the structure of these multi-adjective strings, we treat adjectival modification as syntactic adjunction, as in (1). Semantically, we treat modification as set intersection, where the adjective restricts the set characterized by the nominal denotation to just those elements that hold the specified property. For our purposes, it does not matter whether this intersection proceeds

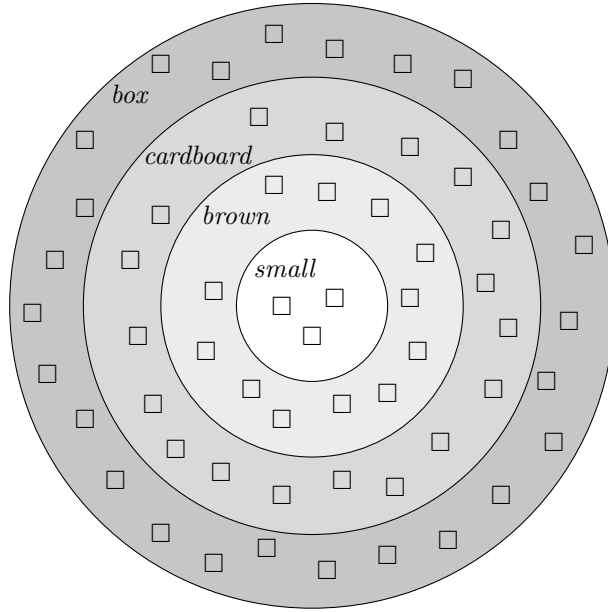
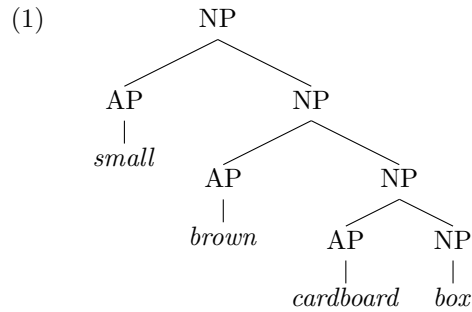


Figure 1: Something.

via a special mode of semantic composition (e.g., Predicate Modification; Heim and Kratzer, 1998), via run-of-the-mill functional application with adjectives of a higher type (Parsons, 1970; Montague, 1970; Kamp, 1975; Siegel, 1986), or via functional structure (Rubin, 1994; Scontras and Nicolae, 2014). In each case, semantic composition proceeds outward from the noun, with adjectives closer to the noun making their semantic contribution earlier than adjectives farther away. The resulting nominal denotation appears in (2), where the NP characterizes the set of small brown cardboard boxes.



(2) $\llbracket \text{small brown cardboard box} \rrbracket =$

$$\lambda x. \text{box}(x) = \text{cardboard}(x) = \text{brown}(x) = \text{small}(x) = 1$$

To see how we arrive at the denotation in (2), consider the illustration of this

process in Figure 1. Each circle corresponds to the denotation of an NP node in (1), with the elements (\square) contained within that circle representing elements of the nominal denotation. The outermost circle represents the denotation of the smallest NP, *box*. In this toy example, there are 59 boxes. Moving inward, we arrive at the next-highest NP denotation, *cardboard box*; there are only 33 such boxes, so the 26 boxes that are not cardboard are pruned from the denotation. Moving inward still, we get the 15 boxes that are both brown and cardboard; the 18 non-brown cardboard boxes have been discarded. Finally, at the innermost circle, we have just those three boxes that are at once cardboard, brown, and small; the 12 brown cardboard boxes that are not small get ignored.

Thinking about the contributions of the adjectives in the example above, we notice that, for the purpose of establishing reference, different adjectives do different amounts of work. Here, “work” gets equated with reference-establishing potential, or information gain. Measured in terms of the number of possible referents ruled out, more work is done by the adjectives closer to the noun. In Figure 1, *cardboard* makes the largest cut in the noun’s denotation, pruning 26 adjectives. Put differently, *cardboard* operates over the largest set: for each of the 59 boxes, one must decide whether or not it is cardboard. Thus, there are 59 opportunities to make a mistake in this decision process, wherein a listener might misjudge a box as cardboard or not. The next adjective, *brown*, makes a smaller cut, operating only over the 33 cardboard boxes; thus, there are fewer possibilities for error. The last adjective, *small*, operates over the smallest set, with the smallest chance of error.

Here is the crux of the account, and finally a return to the issue of subjectivity: less subjective content is more useful for effectively communicating about the world. Encountering a relatively objective adjective like *cardboard*, a listener arrives at a precise concept—one that closely aligns with that of the speaker who uttered the adjective. More subjective adjectives introduce the potential for errors in alignment, as speakers and listeners might (faultlessly) disagree about category boundaries. When it comes to ordering preferences, speakers consolidate the less subjective, more useful content around the modified noun. The claim is that they do so in an attempt to aid the listener in establishing reference by minimizing errors in alignment.

Subjectivity-based adjective ordering preferences emerge once speakers take into account the perspective of their listeners. With the goal of establishing nominal reference, less subjective adjectives are less likely to lead to errors in classification, where a listener could have a diverging opinion about whether or not some objects hold the relevant property. Although in English we encounter the reverse order, modification proceeds outward from the noun. Thus, speakers employ the most useful, least subjective adjectives early in this semantic process where there is the greatest potential for misalignment. The story works the same in languages with post-nominal adjectives where we find mirror-image preferences: linear distance corresponds to hierarchical distance, and adjectives that are closer to the noun make their semantic contributions earlier.

If we are on the right track in assuming that pressures from successful reference resolution, together with awareness of potential disagreement between

speakers and listeners, leads to cross-linguistically robust adjective ordering preferences, the question turns next to how these preferences develop and how they get represented. For now we can only gesture toward possible answers. In a recent corpus analysis of child-directed and child-produced speech, Bar-Sever et al. (2017) documented the emergence of adult-like ordering preferences by the age of four. But are children engaging in the sophisticated theory-of-mind reasoning described above as they form these preferences? Probably not. A growing body of evidence suggests that children struggle with adult-like subjectivity awareness long after ordering preferences emerge (Foushee and Srinivasan, 2017). It would seem, then, that rather than deploying subjectivity-based heuristics, children are merely tracking and reflecting the statistics of their input, a task they are known to excel at (XXX). Children might categorize the regularities of their input according to semantic classes or adjective function, but the ultimate source of these regularities remains the careful interaction of property subjectivity with successful reference resolution.

5. Conclusion

Adjective subjectivity predicts adjective ordering preferences, a remarkably stable property of language design. Here we have attempted to answer the question of why subjectivity should play the role it does in these preferences. Subjective content allows for miscommunication to arise if speakers and listeners arrive at different judgments about a property description. Hence, less subjective content is more useful at communicating about the world. Speakers deploy this more useful content early in the semantic construction of nominals, as reflected in the hierarchical structure of modification: noun phrases are built semantically outward from the noun, and less subjective content enters earlier into this process. This reference-resolution account of subjectivity-based ordering preferences meets the desiderata explored in Section 3: it is predicated upon the findings of Scontras et al. (2017), and so it enjoys firm empirical support. No less important, the current proposal extends seamlessly to cover the mirror-image preferences in post-nominal languages. Perhaps most appealing is the broad applicability of the proposed account: we find the same preferences cross-linguistically because communication is a central goal of language use, so pressures for successful communication apply universally.