On the grammatical source of adjective ordering preferences

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

Word count: 7764

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 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 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 hypothesis 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 naturalistic productions from English corpora. Finding an extremely strong correlation between the behavioral measure and the corpus counts ($r^2 = .83$, 95% CI [.63, .90]), the authors concluded 1) that naïve speakers have reliable and robust adjective ordering preferences, and 2) that the behavioral measure faithfully captured these 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 "subjectiv-

2

ity" 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 ($r^2 = .91$, 95% CI [.86, .94]), Scontras et al. concluded that both measures successfully capture adjective subjectivity.

It bears noting that many factors can contribute to the perceived subjectivity of an adjective, including semantic notions typically thought of as vagueness (e.g., red by which standard?), evaluativity (e.g., beautiful according to whom?), or relativeness/context dependence (e.g., large compared to what?). Moreover, there exist a variety of semantic theories designed to account for these specific notions (e.g., the supervaluations of Kamp and Partee 1995, the perspectives of Kölbel 2002, or the judges of Lasersohn 2005). There are also various formal notions of the subjective vs. objective distinction defined in terms of judges (Sæbø, 2009), counterstances (Kennedy and Willer, 2016), outlooks (Coppock, 2018), etc. For our purposes and the purposes of Scontras et al.'s study, the semantic source of subjectivity runs orthogonal to the simple fact that speakers have stable estimates of subjectivity operationalized via faultless disagreement. In other words, whatever its source, language users recognize that certain adjectives can lead more frequently to cases of misalignment where people might (faultlessly) disagree about the set of things picked out by a given adjective. It is this notion that we refer to with the label "subjectivity."

With clear estimates of adjective subjectivity and of the preferences themselves, 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 inherentness (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 postnominal 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 continued 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 orthogonal 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 effi-

ciently 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 pre-nominal adjective ordering, and it might extend to handle the mirror-image preferences in post-nominal languages—though the details would need to be carefully thought through. 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's 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 in the broad sense—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)
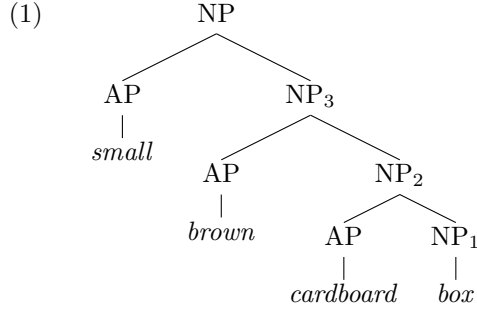
Properties implied by the nominal have meanings that are at least partially contained already within the nominal meaning, hence the implication (cf. the notion of mutual informativity; Futrell, 2017). Determiners that are less implied by the nominal will be more informative (i.e., unexpected) when encountered, 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 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 full NP characterizes the set of small brown cardboard boxes.

(1)

$$
\begin{array}{c}
\text{NP} \\
\diagup \quad \diagdown \\
\text{AP} \qquad \text{NP}_3 \\
| \qquad \diagup \quad \diagdown \\
\textit{small} \quad \text{AP} \qquad \text{NP}_2 \\
\qquad | \qquad \diagup \quad \diagdown \\
\qquad \textit{brown} \quad \text{AP} \quad \text{NP}_1 \\
\qquad \qquad | \qquad | \\
\qquad \qquad \textit{cardboard} \quad \textit{box}
\end{array}
$$

(2) $[\![$small brown cardboard box$]\!] =$

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

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.
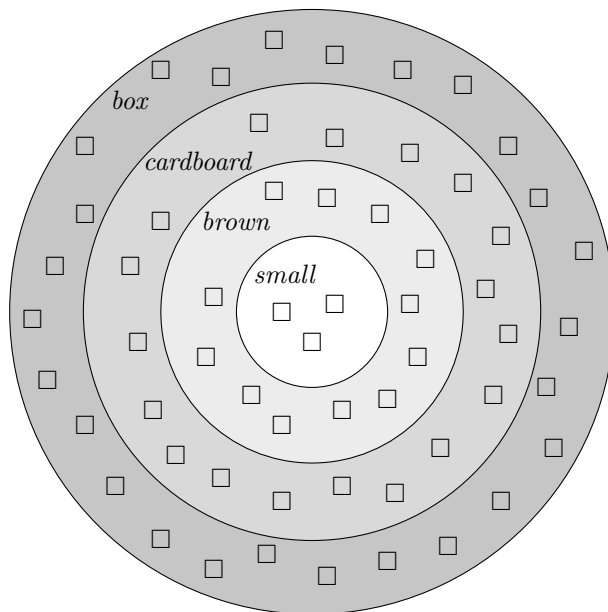
Figure 1: An illustration of restrictive modification in *small brown cardboard box*.

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 potential for information gain. Measured in terms of the number of possible referents considered, 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 (i.e., establishing reference). 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.

The following subsection walks through a concrete example of how ordering with respect to decreasing subjectivity minimizes alignment errors, thereby maximizing the probability of successful referent resolution. Readers already convinced (or wary of math) should skip to the subsection that follows, where we discuss potential worries and further avenues to explore.

4.1. A mathematical demonstration

To get our story off the ground, we must consider the semantics of modification in more detail. To model the potential for faultless disagreement in subjective properties, we introduce noise into the semantics of our adjectives. For each potential referent an adjective classifies, we introduce the potential for misclassification ϵ_{adj}, which stands proxy for the adjective's subjectivity.[1] On the basis of ϵ_{adj}, each adjective has some probability $p_{adj}(\texttt{obj})$ of correctly classifying some object $\texttt{obj}$:

$$p_{adj}(\texttt{obj}) = 1 - \epsilon_{adj} \tag{3}$$

[gcs: adding back in misclassification potential] Stacking a series of noisy adjectives together in a modification structure, there will be a greater chance of incorrect classifications (i.e., misalignments) if the adjectives are not ordered according to subjectivity. As the number of objects to be classified in the nominal denotation (i.e., the cardinality of the NP denotation $|\text{NP}|$) increases, so too does the probability of misclassification $p(\text{error})$:[2]

$$p(\text{error}_{adj}) = 1 - p(\text{no-error}_{adj})$$
$$= 1 - (1 - \epsilon_{adj})^{|\text{NP}|} \tag{4}$$

For a multi-adjective string, we can calculate the probability of *any* errors in misclassification by multiplying the individual error probabilities from each adjective as in (5); crucially, $|\text{NP}|$ will decrease as adjectives restrict the nominal denotation.

---

[1]A general schema for formalizing the potential for misclassification in the semantics of an adjective appears in (i):

(i) $[\![\text{ADJ}]\!] = \lambda x.$ if $\texttt{ADJ}(x)$ then $\texttt{flip}(1 - \epsilon_{adj})$, else $\texttt{flip}(\epsilon_{adj})$

The function $\texttt{flip(x)}$ returns a sample from a Bernoulli distribution, where a random variable takes the value 1 with probability $\texttt{x}$; one can think of this function as simulating the outcome of a weighted coin flip, where heads corresponds to 1 (i.e., true) and tails corresponds to 0 (i.e., false). For our purposes, the addition of $\texttt{flip()}$ to an adjective's semantics introduces noise at the rate ϵ, where ϵ increases with subjectivity.

[2]For convenience, we identify NP with its extension, a set of objects.

$$p(\text{error}_{NP}) = 1 - (1 - \epsilon_{adj_1})^{|\text{NP}_1|} \cdot \ldots \cdot (1 - \epsilon_{adj_n})^{|\text{NP}_n|} \tag{5}$$

Ordering with respect to subjectivity minimizes the probability of misclassifications for a multi-adjective string by ensuring that $|\text{NP}|$ decreases as ϵ increases. For a concrete example, suppose there are three boxes: box_s is small but not brown, box_b is brown but not small, and box_{sb} is both small and brown. To calculate the probability of any misclassifications for the two adjective orderings *small brown* vs. *brown small*, we will need each adjective's potential for misclassification ϵ; we use the subjectivity scores from Scontras et al. (2017) to set these values: 0.20 for *brown* and 0.64 for *small*.[3] As the following calculations demonstrate, ordering with respect to decreasing subjectivity, (6), results in a lower probability for misclassifications than the reverse order, (7):

$$p(\text{error}_{small\ brown\ box}) = 1 - (1 - 0.20)^3 \cdot (1 - 0.64)^2 = \mathbf{0.93} \tag{6}$$

$$p(\text{error}_{brown\ small\ box}) = 1 - (1 - 0.64)^3 \cdot (1 - 0.20)^2 = \mathbf{0.97} \tag{7}$$

[gcs: end of misclassification potential]

However, minimizing misclassifications and correctly classifying the intended referent, thereby allowing for successful reference resolution, are subtly differently notions. If we assume a fixed misclassification potential ϵ for each adjective and a truly intersective semantics for modification, the probability of correctly classifying the intended referent once a noun has been modified by multiple adjectives does not depend on the order of the adjectives (i.e., the order of semantic composition). Given the commutativity of noise in intersective modification, a nominal with two adjectives will correctly classify the intended referent with probability $p_{adj_1}(\texttt{ref}) \cdot p_{adj_2}(\texttt{ref})$. In *small brown box*, the probability that the intended referent will remain in the full nominal denotation is equal to $p_{brown}(\texttt{ref}) \cdot p_{small}(\texttt{ref})$, irrespective of order. If the pressure for subjectivity-based ordering preferences arises out of pressures toward successful reference resolution, the commutativity of noise in intersective modification will not deliver subjectivity-based preferences.[4]

We break the commutativity of noise and thus the order-independence of modification once we recognize that ϵ_{adj} is not a fixed value, but rather varies with the size of the set to be classified, $|\text{NP}|$ (i.e., the number of objects under consideration). Each classification that must be made takes some computational processing. If we posit a fixed processing budget, more classifications

---

[3]These values are adopted for illustrative purposes only. Subjectivity scores are likely an inflated estimate of an adjective's potential for misclassification ϵ. Still, any values for ϵ_{small} and ϵ_{brown} would do so long as $\epsilon_{small} > \epsilon_{brown}$.

[4]Not all modification is intersective (Kamp and Partee, 1995; Truswell, 2009; McNally, 2016). Indeed, intersective modification is likely a poor choice for adjectives like *small*: asserting that Dumbo is a small elephant need not commit the speaker to the assertion that Dumbo is small. Instead, many cases of modification are subsective, such that the interpretation of the modifier (i.e., the adjective) depends crucially on the denotation of its complement (i.e., the nominal it modifies). The account developed below applies just as well to all cases of restrictive modification, including subsective ones.

will necessarily mean making each with fewer resources. Making a stochastic classification with less computation can be done at the expense of precision, that is, with more noise. Classification noise will monotonically increase with the size of the set that must be classified. Thus, as $|\text{NP}|$ increases, the precision of each individual classification decreases and so the potential for misclassification grows. We model this tendency by revising $p_{adj}(\texttt{ref})$ so that ϵ_{adj} depends on the size of the NP denotation that adj restricts, with the constraint that $\epsilon_{adj}(|\text{NP}|) \le \epsilon_{adj}(|\text{NP}| + 1)$:

$$p_{adj}(\texttt{ref}, \text{NP}) = 1 - \epsilon_{adj}(|\text{NP}|) \tag{8}$$

In a case with two adjectives as in (7), the probability that the full multi-adjective NP correctly classifies the intended referent $\texttt{ref}$, $p_{NP}(\texttt{ref})$, takes into account the probability that adj_1 correctly classifies $\texttt{ref}$ in NP_1, $p_{adj_1}(\texttt{ref}, \text{NP}_1)$, and that adj_2 correctly classifies $\texttt{ref}$ in NP_2, $p_{adj_2}(\texttt{ref}, \text{NP}_2)$.[5] However, $p_{adj_2}(\texttt{ref}, \text{NP}_2)$ depends on the size of NP_2, which itself depends on potential classification errors from adj_1. We must therefore consider all possible values for NP_2.

(7)

NP

AP　　NP$_2$

adj_2　AP　NP$_1$

adj_1　$noun$
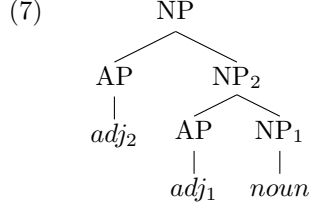
To calculate the probability of each possible NP_2, $P(\texttt{obj} \in \text{NP}_2 | \text{NP}_1, adj_i)$ in (8) looks up the probability that adj_i (in)correctly classifies each potential element $\texttt{obj}$; $V(\texttt{obj}, adj_i)$ serves as our ground truth from the speaker's perspective, returning $\texttt{true}$ just in case $\texttt{obj}$ actually holds the property named by adj_i. The first case in (8) corresponds to the probability of correctly including an element in NP_2 that holds the property named by adj_i; the second case corresponds to the probability of correctly excluding elements that do not hold the relevant property. The final two cases correspond to the probabilities associated with misclassifications.

$$P(\texttt{obj} \in \text{NP}_2 | \text{NP}_1, adj_i) = \begin{cases} 1 - \epsilon_i(|\text{NP}_1|) \text{ if } \texttt{obj} \in \text{NP}_2 \ \& \ V(\texttt{obj}, adj_i) \\ 1 - \epsilon_i(|\text{NP}_1|) \text{ if } \texttt{obj} \notin \text{NP}_2 \ \& \ \neg V(\texttt{obj}, adj_i) \\ \epsilon_i(|\text{NP}_1|) \text{ if } \texttt{obj} \notin \text{NP}_2 \ \& \ V(\texttt{obj}, adj_i) \\ \epsilon_i(|\text{NP}_1|) \text{ if } \texttt{obj} \in \text{NP}_2 \ \& \ \neg V(\texttt{obj}, adj_i) \end{cases} \tag{8}$$

Our aim is the probability of successful referent classification at the level of NP, $p_{NP}(\texttt{ref})$. In other words, we want to know how probable it is that both
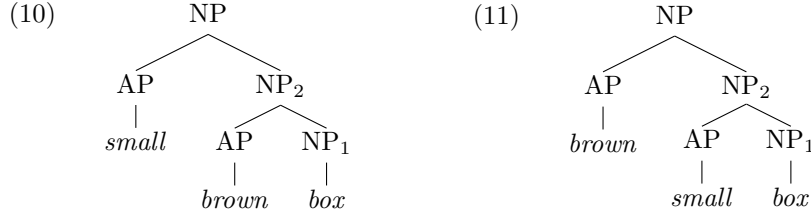
---

[5]For simplicity, we assume a fixed (i.e., noiseless) extension for the noun.

adjectives correctly classify the intended referent. To calculate this probability, we sum over possible values of NP_2 where `ref` was correctly classified by adj_1 (i.e., where $\text{ref} \in NP_2$). For each potential NP_2, we then find the probability of successful classification by adj_2, $p_{adj_2}(\text{ref}, NP_2)$, and multiply it by the probability of having arrived at that NP_2:

$$p_{NP}(\text{ref}) = \sum_{\substack{NP_2 \subseteq NP_1 \\ \text{where } \text{ref} \in NP_2}} p_{adj_2}(\text{ref}, NP_2) \cdot \prod_{x \in NP_1} P(x \in NP_2 | NP_1, adj_1) \quad (9)$$

Adjectives closer to the noun will compose earlier semantically, and so the number of potential referents they must classify will be larger (cf. the example in Fig. 1). In (9), this fact ensures that $|NP_2| \leq |NP_1|$. Because adjectives that compose earlier classify a larger set, we maximize $p_{NP}(\text{ref})$ by ensuring that adjectives with lower subjectivity (i.e., with a lower ϵ) compose earlier.

To see the role of subjectivity-based ordering in maximizing the probability of successful reference resolution, consider the choice between *small brown box* vs. *brown small box* in (10) vs. (11).

(10)

```
            NP
          /    \
        AP      NP₂
        |      /    \
      small  AP     NP₁
             |       |
           brown    box
```

(11)

```
            NP
          /    \
        AP      NP₂
        |      /    \
      brown  AP     NP₁
             |       |
           small    box
```

Suppose once again that there are three boxes: box_s, box_b, and box_{sb}. To calculate the probability of successfully resolving the referent to the small brown box for the two adjective orderings, we will need each adjective's potential for misclassification; here, we use the subjectivity scores from Scontras et al. (2017) to set the upper limit of these values (i.e., when $|NP| = 3$), and assume that ϵ decreases by 0.04 with each decrease in $|NP|$.[6] We adopt the following values:

$$\epsilon_{small}(|NP| = 3) = 0.64 \qquad \epsilon_{brown}(|NP| = 3) = 0.20$$
$$\epsilon_{small}(|NP| = 2) = 0.60 \qquad \epsilon_{brown}(|NP| = 2) = 0.16$$
$$\epsilon_{small}(|NP| = 1) = 0.56 \qquad \epsilon_{brown}(|NP| = 1) = 0.12 \quad (12)$$

In the *small brown box* vs. *brown small box* example, the first adjective to compose will operate over the set of three boxes (i.e., $|NP_1| = 3$). The second adjective will operate over a set that has been restricted by the first adjective, so we ensure that $|NP_2| \leq |NP_1|$. Starting with the preferred order in (10), there are four possibilities for NP_2 (i.e., for ⟦brown box⟧) that include the intended

---

[6] As before, these values are adopted for illustrative purposes only. The decrease of 0.04 for decreasing $|NP|$ is chosen arbitrarily.

referent. In (13), we list the possible extensions of NP_2, together with the probability of each box's classification in parentheses (i.e., $P(\texttt{obj} \in NP_2|NP_1, brown)$ for each box $\texttt{obj}$); (13) also lists the probability of correctly classifying $\texttt{ref}$ in NP_2 ($p_{small}(\texttt{ref}, NP_2)$). Multiplying across the rows, we arrive at the probability p of correctly classifying $\texttt{ref}$ for each possible NP_2; summing over the values of p, we arrive at the probability of correctly classifying $\texttt{ref}$ for the full NP: 0.32.

$[\![$brown box$]\!]$				
box_{sb}	box_b	box_s	$p_{small}(\texttt{ref}, NP_2)$	p
T (0.80)	T (0.80)	T (0.20)	0.36	0.05
T (0.80)	T (0.80)	F (0.80)	0.40	0.20
T (0.80)	F (0.20)	T (0.20)	0.40	0.01
T (0.80)	F (0.20)	F (0.80)	0.44	0.06

$$p_{small\ brown\ box}(\texttt{ref}) = \mathbf{0.32}$$

(13)

In (14), we perform the same calculations for the dispreferred ordering in (11); with this ordering, the probability of correctly classifying $\texttt{ref}$ for the full NP is 0.30.

$[\![$small box$]\!]$				
box_{sb}	box_b	box_s	$1 - \epsilon_{brown}(|NP_2|)$	p
T (0.36)	T (0.64)	T (0.36)	0.80	0.07
T (0.36)	T (0.64)	F (0.64)	0.84	0.12
T (0.36)	F (0.36)	T (0.36)	0.84	0.04
T (0.36)	F (0.36)	F (0.64)	0.88	0.07

$$p_{brown\ small\ box}(\texttt{ref}) = \mathbf{0.30}$$

(14)

As the calculations above demonstrate, ordering with respect to decreasing subjectivity results in a higher probability of successfully classifying the intended referent than the reverse order. Because it is necessarily the case that $|NP_2| \leq |NP_1|$ in the presence of restrictive adjectival modification, this pattern holds generally. [gcs: adding a discussion of the generality] We systematically explored the generality of this pattern with a search through the possible parameter space. By varying ϵ_{small}, ϵ_{brown}, and the size and makeup of the initial nominal denotation (i.e., the number of boxes and their properties), we tested 103,740 cases of multi-adjective modification.[7] Of those cases tested, 93% were such that ordering with respect to decreasing subjectivity resulted in a higher probability of correctly classifying the intended referent. Those cases where the reverse ordering resulted in a higher probability of correct referent classification fell into one of two classes: 1) there was a larger number of objects from the less

---

[7]Minimum values for the noise parameters varied between 0.01 and 0.77 in steps of 0.04, with the constraint that the minimum value of ϵ_{small} exceeded the minimum value of ϵ_{brown}. The initial nominal denotation varied in cardinality between two and five, with the constraint that only one box was both small and brown (i.e., the intended referent); otherwise, boxes took on all possible combinations of the properties of being small and being brown.

subjective category (e.g., one small brown box, two brown boxes, and one small box), $\epsilon_{\mathrm{brown}}$ was sufficiently small, and the difference in ϵ between the two adjectives was sufficiently small; or 2) there was a larger number of objects from the more subjective category (e.g., one small brown box, one brown box, and two small boxes), $\epsilon_{\mathrm{brown}}$ was sufficiently large, and the difference in ϵ between the two adjectives was sufficiently small. Both classes of cases rely on errors in classification by the more subjective adjective to shrink the set of objects that will get classified by the less subjective adjective, resulting in a higher overall probability of correctly classifying the intended referent.

Although there are cases where ordering with respect to decreasing subjectivity does not maximize the probability of correctly classifying the intended referent, the vast majority of cases do follow this trend. It comes as no surprise, then, that language has regularized this strong trend as stable subjectivity-based ordering preferences. We thus see how 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. A simple policy of misclassification avoidance delivers subjectivity-based ordering preference. Subjectivity-based ordering preferences likewise maximize the probability of successful classification of the intended referent in cases of resource-bounded computation. Although in English we encounter the reverse order, modification proceeds semantically 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 proposed account 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.

4.2. Some potential worries

The astute reader will recognize that the proposed account of adjective order, which relies on incremental semantic composition, ostensibly stands at odds with the linear nature of sentence processing, specifically with respect to reference resolution. Eberhard et al. (1995) report the results of a visual-world eye-tracking study featuring nominals with multi-adjective strings; their results suggest that listeners use information from incoming words to prune the set of potential referents *as that information becomes available.* Sedivy et al. (1999) follow up on this finding by demonstrating incremental reference resolution even for non-intersective (i.e., context-dependent) adjectives. The empirical picture appears clear: listeners' eye movements narrow in on potential nominal referents as time progresses linearly.[8] And yet our proposed account assumes that

---

[8]While eye movements might narrow in on the potential nominal referent in visual-world eye-tracking studies, it remains unclear whether the listener's beliefs are similarly narrowed. Recent results from Qing et al. (2018) suggest that eye movements in reference tasks might

semantic composition proceeds outward from the noun, a direction opposite to the linear uptake of words, at least in pre-nominal languages like English.

The pressures that deliver adjective ordering preferences evidence a case where hierarchical, compositional structure appears to take precedence over linear, incremental processing. The work on predictive looks during incremental processing only serves to increase the interest of this tension. However, the early uptake of semantic information evidenced by predictive looking in eye-tracking studies does not rule out that the semantic composition of nominal phrases proceeds outward from the noun; it is this semantic composition process that stands to explain the role of subjectivity in adjective ordering preferences.

A note of caution is in order: this account relies crucially on the assumption that a speaker's goal when using modificational content is to establish reference, yet this might not always be the case. Consider a dialog in which Alex says to Chris, "Do you see what he's wearing?" Chris responds, "What a tacky polyester shirt!" Here, the referent (i.e., the relevant shirt) is already in common ground before Chris's utterance; the adjectives *tacky* and *polyester* are thus unlikely to be employed in the service of establishing reference. Instead, these kinds of non-restrictive uses communicate the speaker's stance toward the shirt in question. It remains an empirical question whether speaker goals influence ordering preferences such that subjectivity plays a lesser role in the absence of reference resolution. It might also be the case that an account of non-restrictive adjective use makes the same predictions regarding the role of subjectivity (cf. Hahn et al., 2018).

Nevertheless, 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, lead 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 abstract knowledge of 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 (e.g., Saffran et al., 1996). 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.

---

be only loosely correlated with the degree to which an object is believed to be the intended referent.

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.

References

Bar-Sever, G., G. Scontras, and L. S. Pearl (2017). Subjective little learners: Hyperarticulated input and the early development of adjective ordering preferences. Boston University Conference on Language Development 42, Boston, MA.

Barker, C. (2013). Negotiating Taste. *Inquiry 56*(2-3), 240–257.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In R. Hayes (Ed.), *Cognition and Language Development*, pp. 279–362. New York: Wiley & Sons, Inc.

Biber, D., S. Johansson, G. Leech, S. Conrad, and E. Finegan (1999). *Longman Grammar of Spoken and Written English*. London: Longman.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formation. *Psychological Review 89*, 1–47.

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

Cinque, G. (1994). On the evidence for partial N-movement in the romance DP. In R. S. Kayne, G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, and R. Zanuttini (Eds.), *Paths Towards Universal Grammar. Studies in Honor of Richard S. Kayne*, pp. 85–110. Washington DC: Georgetown University Press.

Coppock, E. (2018). Outlook-based semantics. *Linguistics and Philosophy 41*(2), 125–164.

Dixon, R. (1982). *Where have all the adjectives gone? And other essays in semantics and syntax.* Berlin: Mouton.

Eberhard, K. M., M. J. Spivey-Knowlton, J. C. Sedivy, and M. K. Tanenhaus (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research 24*(6), 409–436.

Foushee, R. and M. Srinivasan (2017). Could both be right? Children's and adult's sensitivity to subjectivity in langauge. In *Proceedings of the Thirty-Ninth Annual Meeting of the Cognitive Science Society*, London. Cognitive Science Society.

Futrell, R. L. J. (2017). *Memory and Locality in Natural Language.* Ph. D. thesis, Massachusetts Institute of Technologys, Cambridge, MA.

Hahn, M., J. Degen, N. Goodman, D. Jurafsky, and R. Futrell (2018). An information-theoretic explanation of adjective ordering preferences. In *Proceedings of the Fortieth Annual Conference of the Cognitive Science Society.*

Heim, I. and A. Kratzer (1998). *Semantics in Generative Grammar.* Oxford: Blackwell.

Hetzron, R. (1978). On the relative order of adjectives. In H. Seller (Ed.), *Language Universals*, pp. 165–184. Tübingen: Narr.

Hill, F. (2012). Beauty Before Age? Applying Subjectivity to Automatic English Adjective Ordering. In *Proceedings of the NAACL HLT 2012 Student Research Workshop*, pp. 11–16.

Kamp, H. (1975). Two theories about adjectives. In E. Keenan (Ed.), *Formal Semantics of Natural Languages.* Cambridge: Cambridge University Press.

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

Kennedy, C. (2013). Two sources of subjectivity: Qualitative assessment and dimensional uncertainty. *Inquiry 56*(2-3), 258–277.

Kennedy, C. and M. Willer (2016). Subjective attitudes and counterstance contingency. In *Proceedings of SALT 26*, pp. 913–933.

Kölbel, M. (2002). *Truth Without Objectivity.* London and New York: Routledge.

Kölbel, M. (2004). Faultless Disagreement. *Proceedings of the Aristotelian Society 104*, 53–73.

Laenzlinger, C. (2005). French adjective ordering: perspectives on DP-internal movement types. *Lingua 115*, 645–689.

LaPolla, R. J. and C. Huang (2004). Adjectives in Qiang. In R. M. W. Dixon and A. Y. Aikenvald (Eds.), *Adjective Classes: A Cross-Linguistic Typology*, pp. 306–322. Oxford: Oxford University Press.

Lasersohn, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy 28*, 643–686.

MacFarlane, J. (2014). *Assessment Sensitivity*. Oxford: Clarendon Press.

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

Martin, J. E. (1969b). Some competence-process relationships in noun phrases with prenominal and postnominal adjectives. *Journal of Verbal Learning and Verbal Behavior 8*, 471–480.

McNally, L. (2016). Modification. In M. Aloni and P. Dekker (Eds.), *Cambridge Handbook of Formal Semantics*, pp. 442–466. Cambridge: Cambridge University Press.

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

Montague, R. (1970). English as a formal language. In B. Visentini (Ed.), *Linguagi Nella Società e Nella Tecnica*, pp. 189–224. Milan: Edizioni di Comunità.

Parsons, T. (1970). Some problems concerning the logic of grammatical modifiers. *Synthese 21*(3-4), 320–324.

Posner, R. (1986). Iconicity in syntax: The natural order of attributes. In P. Bouissac, M. Herzfeld, and R. Posner (Eds.), *Iconicity: Essays on the Nature of Culture*, pp. 305–337. Tübingen: Stauffenburg Verlag.

Qing, C., D. Lassiter, and J. Degen (2018). What do eye movements in the visual world reflect? a case study from adjectives. In *Submitted to CogSci 2018*.

Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik (1985). A Comprehensive Grammar of the English Language.

Rubin, E. J. (1994). *Modification: A Syntactic Analysis and Its Consequences*. Ph. D. thesis, Cornell University, Ithaca, NY.

Sæbø, K. J. (2009). Judgment ascriptions. *Linguistics and Philosophy 32*(4), 327–352.

Saffran, J. R., R. N. Aslin, and E. Newport (1996). Statistical learning by 8-month-old infants. *Science 274*(5294), 1926–1928.

Scontras, G., J. Degen, and N. D. Goodman (2017). Subjectivity predicts adjective ordering preferences. *Open Mind: Discoveries in Cognitive Science 1*(1), 53–65.

Scontras, G. and A. C. Nicolae (2014). Saturating syntax: Linkers and modification in Tagalog. *Lingua 149*, 17–33.

Scott, G.-J. (2002). Stacked adjectival modification and the structure of nominal phrases. In G. Cinque (Ed.), *The cartography of syntactic structures, Volume 1: Functional structure in the DP and IP*, pp. 91–120. Oxford: Oxford University Press.

Sedivy, J. C., M. K. Tanenhaus, C. G. Chambers, and G. N. Carlson (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition 71*, 109–147.

Seiler, H. (1978). Determination: A functional dimension for interlanguage comparison. In H. Seiler (Ed.), *Language Universals, Papers from the Conference held at Gummersbach/Cologne, Germany, October 3-8, 1976*, pp. 301–328. Tübingen: Narr.

Siegel, M. E. A. (1986). *Capturing the Adjective*. Ph. D. thesis, University of Massachusetts at Amherst, Amherst, MA.

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.

Tucker, G. (1998). *The Lexicogrammar of Adjectives: A Systemic Functional Approach to Lexis*. London: Cassell Academic.

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.

Ziff, P. (1960). *Semantic Analysis*. Ithaca, NY: Cornell University Press.