On the role of conjunction in adjective ordering preferences
Cesar Manuel Rosales Jr. & Gregory Scontras*

Abstract. Adjective ordering preferences are robustly attested in English and many unrelated languages. In nominals with multi-adjective strings (e.g., big blue box), chances are the order of the adjectives is non-arbitrary. However, ordering preferences are claimed to neutralize in cases where multi-adjective strings are formed via conjunction (e.g., blue and big box). We provide empirical evidence in support of this claim, but with an important caveat: conjunction neutralizes adjective ordering preferences in languages like Spanish where multi-adjective strings obligatorily feature conjunction. In English, where multi-adjective strings optionally feature conjunction, ordering preferences persist in the presence of conjunction.

Keywords. adjective ordering; subjectivity; conjunction; Spanish

1. Introduction. In cases of adjectival modification, speakers are not limited to using a single adjective. We may talk of big blue boxes or new plastic chairs. However, the relative order of the two adjectives is not arbitrary: our ordering preferences dictate that big blue box sounds vastly more natural than the reverse order, blue big box. Curiously, the exact same preferences have been reported in a range of diverse languages, both with pre- and post-nominal adjectives (e.g., Hungarian, Dutch, Mandarin Chinese, Indonesian, Tagalog; Martin 1969; Hetzron 1978; Dixon 1982; Sproat & Shih 1991; Samonte & Scontras 2019). It would seem that these preferences attend to the relative distance of adjectives from the nouns they modify.

Recently, Scontras et al. (2017) used behavioral and corpus evidence to document the predictive power of adjective subjectivity in adjective ordering preferences. The authors showed that adjective order in multi-adjective strings is best predicted by the subjectivity of the adjectives involved, such that less subjective adjectives are preferred closer to the modified noun—at least in English. To the extent that other languages demonstrate the English preferences, subjectivity is likely to play a role there as well.

Recent accounts of the role of subjectivity in adjective ordering preferences demonstrate how subjectivity-based orderings maximize communicative success in the incremental construction of a noun phrase. However, these accounts split on whether they assume hierarchical semantic composition (Simonić 2018; Scontras et al. to appear) or a linear, flat process (Hahn et al. 2018). The current work explores a prediction on which these two accounts differ: conjunction disrupts the incremental hierarchical composition of multi-adjective strings, but not its linear ordering; to the extent that conjunction neutralizes ordering preferences, we will have evidence in favor of the hierarchical account. To test this prediction, we explore ordering preference in Spanish, a language with post-nominal adjectives that require conjunction in multi-adjective strings.

The paper is structured as follows. In Section 2, we survey the empirical foundations of subjectivity-based ordering preferences established by Scontras et al. (2017), as well as recent proposals accounting for the role of subjectivity. In Section 3, we extend the original English

*Authors: Cesar Manuel Rosales Jr., University of California, Irvine (cesarmr@uci.edu), Gregory Scontras, University of California, Irvine (g.scontras@uci.edu).
methodology from Scontras et al. (2017) to determine the status of subjectivity-based ordering preferences in Spanish. We return to English in Section 4, where we test for preferences both with and without conjunction. Section 5 concludes with a discussion of our findings.

2. Background. Adjective ordering preferences have received a good deal of attention in both the linguistics and psychology literatures, and with good reason: that we should find the same preferences in such a diverse range of languages must evidence deep properties of language and the minds from which it emerges. Proposals split on whether they attempt to characterize the linguistic representation of these preferences (e.g., Dixon 1982; Cinque 1994; Scott 2002; McNally & Boleda 2004; Laenzlinger 2005; Truswell 2009) or the psychological factors that predict them (e.g., Sweet 1898; Whorf 1945; Ziff 1960; Martin 1969; Seiler 1978). Inspired by accounts linking ordering preferences to properties of adjective meaning, Scontras et al. (2017) pursued the hypothesis that adjective subjectivity predicts adjective ordering preferences (Quirk et al. 1985; Hetzron 1978; Tucker 1998; Hill 2012).

To test the subjectivity hypothesis, Scontras et al. (2017) first established a behavioral measure of English ordering preferences. They presented experimental participants with pairs of multi-adjective strings that differed in the relative order of the adjectives, and asked participants to indicate which order sounded more natural. The authors validated their behavioral measure by comparing it with naturalistic productions from corpora. Next, Scontras et al. established a behavioral measure of adjective subjectivity, which they validated with a faultless disagreement task (Kölbel 2004; Barker 2013; Kennedy 2013; MacFarlane 2014). When comparing their estimates of the ordering preferences with the adjective subjectivity measure, Scontras et al. (2017) found that subjectivity accounted for 85% of the variance in the ordering preferences for 26 adjectives from seven different semantic classes (e.g., size, color, shape, etc.). Looking at a larger collection of 74 adjectives from 13 semantic classes, the authors found that subjectivity accounted for 61% of the variance in ordering preferences.

With strong evidence for the role of subjectivity in adjective ordering, the task shifts next to explaining why subjectivity should play the role it does in the observed preferences. All of the explanations offered to date agree that subjectivity-based orderings maximize communicative success. However, the details of these explanations differ, as does the evidence used to support them.

Simonič (2018) developed a probabilistic model of utterance selection conceived within the Rational Speech Act modeling framework (Franke & Jäger 2016; Goodman & Frank 2016). The model chooses between multi-adjective strings that differ in the relative order of the adjectives (e.g., *big blue box* vs. *blue big box*). To model adjective subjectivity, Simonič assumed that a listener’s semantic entry for an adjective could deviate from that of the speaker; the probability of this deviation increases with subjectivity. Simonič further assumed that meaning composition follows the hierarchical structure of modification, as in (1), such that adjectives that occur farther from the modified noun operate over a restricted set of potential references (e.g., the set of blue boxes vs. the set of boxes).

(1) \[ \text{big} \quad \text{blue} \quad \text{box} \]
With the goal of establishing reference—in other words, signaling a specific object to a listener—Simonič used his utterance choice model to demonstrate how ordering adjectives with respect to decreasing subjectivity is more useful in most cases, and thus more likely to be chosen by the speaker. Communicative success is maximized by ensuring that higher-subjectivity adjectives, whose semantics might vary between speakers and listeners and therefore lead to errors in alignment, operate over the smallest set of potential referents, thereby minimizing opportunities for errors in alignment between speakers and listeners.

Scontras et al. (to appear) offered a similar explanation for subjectivity-based ordering preferences. They modeled adjective subjectivity as the probability that a listener would incorrectly classify a possible referent (e.g., incorrectly classifying some box as blue); more subjective adjectives are more likely to lead to incorrect classifications. Scontras et al. further assumed that classification errors increase with the size of the set to be classified. In other words, classification precision decreases as the number of classifications increases. As with Simonič, these assumptions lead to a higher probability of correct classifications if the most error-prone (i.e., most subjective) adjectives operate over the smallest set of potential referents, composing last in the hierarchical nominal structure. Scontras et al. explored over 100,000 cases of multi-adjective modification, finding that subjectivity-based ordering leads to a higher probability of correct referent classification in 93% of those cases.

Whereas Simonič (2018) and Scontras et al. (to appear) assume hierarchical meaning composition, Hahn et al. (2018) pursued a different approach, assuming instead that listeners incrementally update their beliefs with the linear flow of words. Their model further differs from those of Simonič and Scontras et al. by defining speaker utility not in terms of referential success, but rather in terms of successfully communicating subjective opinions about objects in the world. In other words, Hahn et al. model cases of non-restrictive adjectival modification. They also treat adjective subjectivity as the potential for disagreement between speakers and listeners. However, Hahn et al. introduce memory considerations into their model: words further in the past are more likely to be forgotten. By combining incremental linear processing with memory limitations, the authors predict that speakers will maximize their utility (i.e., maximize the probability of successfully communicating their subjective opinions) by ordering adjectives with respect to decreasing subjectivity.

To summarize, all three models demonstrate how subjectivity-based orderings maximize communicative success. The models differ in the precise details of implementing subjectivity, as well as processing considerations. For present purposes, the most salient difference among these approaches concerns the mode by which meaning gets composed. Simonič (2018) and Scontras et al. (to appear) assume hierarchical meaning composition that proceeds outward from the modified noun; subjectivity-based orderings ensure that adjectives with higher subjectivity compose later, thereby operating over a restricted set of potential referents. On the other hand, Hahn et al. (2018) assume a linear process of meaning composition; subjectivity-based orderings ensure that the most subjective adjectives are most likely to be forgotten. Both styles of approach successfully predict subjectivity-based ordering preferences, and in principle, both sorts of pressures—hierarchical and linear—could be active in the construction of multi-adjective strings. However, the two styles of approach make different predictions when it
comes to conjunction.

Conjunction has been claimed to neutralize ordering preferences in English (e.g., Ford & Olson 1975; Byrne 1979). In effect, this neutralization leads to the acceptability of otherwise-unacceptable adjective orderings. For example, whereas blue big box is strongly dispreferred, blue and big box should be acceptable. When we consider why pressures for subjectivity-based orderings should dissipate in the context of conjunction, the hierarchical accounts of Simonič (2018) and Scontras et al. (to appear) offer a ready explanation. To reiterate the proposed mechanism: subjectivity-based orderings maximize communicative success by ensuring that the most subjective (i.e., noisiest, least useful) adjectives compost last, operating over the smallest set of potential referents. In cases of conjunction, as schematized in (2), this situation no longer holds.

(2)

\[
\text{big and blue box}
\]

With conjunction, the adjectives no longer compose incrementally with the nominal; instead, the adjectives first conjoin together, and then the conjoined structure jointly modifies the noun. Without incremental restriction, it is no longer the case that adjectives farther from the noun will operate over a restricted set of potential referents. Thus, without incremental restriction, it is no longer possible to reduce errors in alignment between speakers and listeners by placing more subjective adjectives farther from the modified noun. Crucially, the linear processing account of Hahn et al. (2018) makes a different prediction: adjectives farther from the noun are still more likely to be forgotten, so pressures toward subjectivity-based ordering should apply with conjunction as they do without it.

We therefore set out to test the predictions of linear vs. hierarchical accounts of subjectivity-based ordering preferences. To do so, we first look at ordering preferences in Spanish, were multi-adjective strings require conjunction. We then turn to English, bringing behavioral data to bear on the status of adjective ordering preferences with conjunction.

3. Experiment 1: Spanish. Spanish offers a natural testbed for teasing apart the predictions of linear vs. hierarchical accounts of subjectivity-based ordering preferences. To do so, we first look at ordering preferences in Spanish, were multi-adjective strings require conjunction. We then turn to English, bringing behavioral data to bear on the status of adjective ordering preferences with conjunction.

3.1 ORDERING PREFERENCES. First, we must determine the status of Spanish ordering preferences. To do so, we replicated the methodology of Expt. 1: Ordering preferences from Scontras et al. (2017) using Spanish materials.1

(3) la caja grande y azul
   the box big and blue
   ‘the big and blue box’

To determine the status of ordering preferences in Spanish, we replicated the methodology from Scontras et al. (2017) using Spanish translations of the English materials.

1 Some adjectives in Spanish may occur pre-nominally. However, this strategy is not fully productive in the language. We therefore focus only on post-nominal multi-adjective strings.
Participants. We recruited 224 participants through Amazon.com’s Mechanical Turk crowdsourcing service. All participants were compensated for their participation. On the basis of their responses to a post-test demographics questionnaire, we identified 48 native speakers of Spanish (30 female, 18 male; mean age: 34). These participants described themselves as having learned Spanish as their first language, and still using Spanish as their primary language at the time of testing; they also reported living in a Spanish-speaking country both before and after the age of eight. Their data were included in the analyses reported below.

Procedure. Participants were tasked with providing judgments about the naturalness of object descriptions. Participants completed a total of 26 trials. On each trial, they indicated which of two object descriptions sounded more natural. To provide their ratings, participants adjusted a slider with endpoints labeled with the competing object descriptions. Each description featured two adjectives, and the description pairs differed only in the relative order of the two adjectives (e.g., la caja grande y azul vs. la caja azul y grande). Adjectives and nouns were chosen at random from the materials in Table 1, with the constraint that each description contained adjectives from two different semantic classes. Where possible, these materials were direct translations of the English materials from Scontras et al. (2017); in place of the English material adjectives (i.e., metal, plastic, wooden), we substituted three nationality adjectives (i.e.,

<table>
<thead>
<tr>
<th>adjective</th>
<th>translation</th>
<th>class</th>
<th>adjective</th>
<th>translation</th>
<th>class</th>
</tr>
</thead>
<tbody>
<tr>
<td>fresco</td>
<td>fresh</td>
<td>age</td>
<td>bueno</td>
<td>good</td>
<td>quality</td>
</tr>
<tr>
<td>nuevo</td>
<td>new</td>
<td>age</td>
<td>malo</td>
<td>bad</td>
<td>quality</td>
</tr>
<tr>
<td>podrido</td>
<td>rotten</td>
<td>age</td>
<td>cuadrado</td>
<td>square</td>
<td>shape</td>
</tr>
<tr>
<td>viejo</td>
<td>old</td>
<td>age</td>
<td>redondo</td>
<td>round</td>
<td>shape</td>
</tr>
<tr>
<td>amarillo</td>
<td>yellow</td>
<td>color</td>
<td>corto</td>
<td>short</td>
<td>size</td>
</tr>
<tr>
<td>azul</td>
<td>blue</td>
<td>color</td>
<td>enorme</td>
<td>huge</td>
<td>size</td>
</tr>
<tr>
<td>marrón</td>
<td>brown</td>
<td>color</td>
<td>grande</td>
<td>big</td>
<td>size</td>
</tr>
<tr>
<td>morado</td>
<td>purple</td>
<td>color</td>
<td>largo</td>
<td>long</td>
<td>size</td>
</tr>
<tr>
<td>rojo</td>
<td>red</td>
<td>color</td>
<td>minúsculo</td>
<td>tiny</td>
<td>size</td>
</tr>
<tr>
<td>verde</td>
<td>green</td>
<td>color</td>
<td>pequeño</td>
<td>small</td>
<td>size</td>
</tr>
<tr>
<td>americano</td>
<td>Mexican</td>
<td>nationality</td>
<td>duro</td>
<td>hard</td>
<td>texture</td>
</tr>
<tr>
<td>español</td>
<td>Spanish</td>
<td>nationality</td>
<td>liso</td>
<td>smooth</td>
<td>texture</td>
</tr>
<tr>
<td>mexicano</td>
<td>American</td>
<td>nationality</td>
<td>suave</td>
<td>soft</td>
<td>texture</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>noun</th>
<th>translation</th>
<th>class</th>
<th>noun</th>
<th>translation</th>
<th>class</th>
</tr>
</thead>
<tbody>
<tr>
<td>manzana</td>
<td>apple</td>
<td>food</td>
<td>escritorio</td>
<td>desk</td>
<td>furniture</td>
</tr>
<tr>
<td>plátano</td>
<td>banana</td>
<td>food</td>
<td>silla</td>
<td>chair</td>
<td>furniture</td>
</tr>
<tr>
<td>queso</td>
<td>cheese</td>
<td>food</td>
<td>sofá</td>
<td>couch</td>
<td>furniture</td>
</tr>
<tr>
<td>tomate</td>
<td>tomato</td>
<td>food</td>
<td>televisión</td>
<td>TV</td>
<td>furniture</td>
</tr>
<tr>
<td>zanahoria</td>
<td>carrot</td>
<td>food</td>
<td>ventilador</td>
<td>fan</td>
<td>furniture</td>
</tr>
</tbody>
</table>

Table 1: Spanish adjectives and nouns used in Experiment 1. Words in italics indicate cases where direct translation from the original English materials was not possible.
Figure 1: Naturalness ratings from Experiment 1 grouped by adjective semantic class. Higher values indicate that a class’s adjectives are preferred farther from the modified noun; lower values indicate that a class’s adjectives are preferred closer. The dashed line indicates chance level, or the absence of stable preferences. Error bars represent bootstrapped 95% confidence intervals drawn from 10,000 samples of the data. Red dots represent point estimates for individual adjectives.

Results. For each of the 26 adjectives in Table 1, we computed a mean preferred-distance measure (i.e., how far the adjective was preferred from the modified noun) by averaging across participants. Figure 1 plots these preferred distance measures grouped by adjective semantic class. There, we see that in all but one of the semantic classes (i.e., quality), participants fail to provide systematic ratings that would evidence stable ordering preferences (i.e., a preference to place adjectives closer or farther from the modified noun); we find a similar pattern of responses at the level of individual adjectives.

3.2 Subjectivity. Having documented the (absence of) ordering preferences in Spanish, we turn now to measuring Spanish adjective subjectivity. Perhaps we fail to find stable ordering preference in Spanish because participants perceive the subjectivity of the adjectives as similar. In other words, perhaps subjectivity predicts ordering preferences in Spanish as it does in English, but what differs between the two languages is the perceived subjectivity of the adjectives involved. To measure subjectivity, we replicated the methodology of Expt. 1: Faultless disagreement validation from Scontras et al. (2017).

Participants. We recruited 106 participants who did not take part in the ordering preferences experiment through Amazon.com’s Mechanical Turk. All participants were compensated for their participation. Using the same exclusion criteria from the ordering preference experiment, we identified 21 native speakers of Spanish (9 female, 12 male; mean age: 33); their data were included in the analyses reported below.
Figure 2: Faultless disagreement ratings from Experiment 1 grouped by adjective semantic class. Higher values indicate that a class’s adjectives are more subjective. Error bars represent bootstrapped 95% confidence intervals drawn from 10,000 samples of the data. Red dots represent point estimates for individual adjectives.

Procedure. Participants read a series of short dialogs in which two speakers disagreed about a property description for some object. For example, one speaker might assert that a box is big, while the other would disagree, saying that the box is not big. Participants judged whether both speakers could be right in their diverging statements; they adjusted a slider with endpoints labeled ‘no, somebody was wrong’ (coded as 0) and ‘yes, the two may be right’ (coded as 1). Participants completed a total of 26 trials, one for each of the adjectives in Table 1.

Results. For each adjective, we computed a mean perceived subjectivity score by averaging the faultless disagreement ratings across participants. Figure 2 plots perceived subjectivity scores grouped by adjective class. Unlike with the ordering preferences, here we see that classes deviate from each other with respect to subjectivity. Below, we will use these scores to test the predictive power of subjectivity in the Spanish ordering preferences.

3.3 Comparing Ordering Preferences with Subjectivity. Finally, we compare the Spanish ordering preferences with perceived subjectivity to evaluate the extent to which subjectivity predicts adjective ordering preferences in Spanish. Figure 3 plots the ordering preferences against the subjectivity scores (i.e., faultless disagreement ratings) for each of the 26 adjectives. Adjective subjectivity accounts for 0.01% of the variance in the ordering preferences ($r^2 = 0.01; 95\%\ CI [0.00, 0.06]$). In other words, subjectivity does not predict ordering preferences in Spanish. However, this finding is not a failure of subjectivity, but rather a byproduct of the absence of Spanish ordering preferences. As we would expect if ordering preferences arise from the hierarchical structure of nominal modification, when conjunction interrupts the incremental adjective composition, as in Spanish, ordering preferences do not arise. We follow up on this finding in the next section by testing the effect of conjunction on English ordering preferences.
4. Experiment 2: English. Having failed to find ordering preferences in the presence of conjunction in Spanish, we turn next to English, where multi-adjective strings optionally feature conjunction. If we are on the right track in assuming that the pressure for subjectivity-based ordering preferences does not arise when conjunction interrupts the incremental semantic composition of multi-adjective modification, then we should find that conjunction neutralizes ordering preferences in Spanish, as it does in English. We therefore replicated the methodology of Expt. 1: Ordering preferences from Scontras et al. (2017), adding in a manipulation whereby some multi-adjective strings featured conjunction.

4.1 Participants. We recruited 60 participants through Amazon.com’s Mechanical Turk. All participants were compensated for their participation. 59 participants identified as native speakers of English (41 female, 18 male; mean age: 36); their data were included in the analyses reported below.

4.2 Procedure. The design of this experiment was similar to that of Expt. 1: Ordering preferences from Scontras et al. (2017), with one crucial difference: the pairs of multi-adjective object descriptions that participants encountered appeared either with or without conjunction. For example, participants might judge the naturalness of big blue box vs. blue big box (i.e., without conjunction), or they might see big and blue box vs. blue and big box (i.e., with conjunction). Participants completed a total of 26 trials. For each object description, adjectives and nouns were chosen at random and came from the original English materials from Scontras et al. (2017); whether or not the description featured conjunction was chosen at random on each trial.

4.3 Results. As with the Spanish ordering preferences, we calculated a mean preferred-distance measure both with and without conjunction by averaging across participants. Figure 4 plots these preferred distance measures against the English adjective subjectivity scores from Scontras et al. (2017). As expected, subjectivity accounts for nearly all of the variance.

Figure 3: Spanish ordering preferences plotted against subjectivity scores for each of the 26 adjectives tested.
in multi-adjective strings without conjunction ($r^2 = 0.89$; 95% CI [0.81, 0.94])—this result replicates the original findings of Scontras et al. (2017). What comes as a surprise, however, is the presence of subjectivity-based ordering preferences with conjunction: although the effect is weaker with conjunction, subjectivity continues to predict ordering preferences ($r^2 = 0.68$; 95% CI [0.40, 0.83]). This finding stands in stark contrast to the Spanish results, where subjectivity accounted for next to no variance in the ordering preferences. In other words, in the presence of conjunction, subjectivity-based ordering preferences weaken but persist in English.

5. General discussion. We have found mixed results on the status of subjectivity-based ordering preferences in the presence of conjunction. In Spanish, where multi-adjective strings are post-nominal and require conjunction, we failed to find evidence of stable ordering preferences, subjectivity-based or otherwise. In English, where multi-adjective strings optionally feature conjunction, we found that with conjunction subjectivity-based ordering preferences weaken but persist. This result is surprising in light of previous claims regarding the role of conjunction in English ordering preferences. Returning to the explanations for subjectivity-based preferences discussed in Section 2, the hierarchical accounts of Simonič (2018) and Scontras et al. (to appear) straightforwardly predict the Spanish findings: conjunction interrupts the incremental hierarchical composition, such that adjectives farther from the noun no longer classify a smaller set out potential referents. Without incremental restriction, the pressures for subjectivity-based orderings are no longer present, thereby accounting for the lack of preferences in Spanish.

However, the story we have just told applies also to cases of conjunction in English, but in English we continue to find stable subjectivity-based preferences in the presence of conjunction. Either we are wrong about our explanation for the absence of stable preference in Spanish, or some other factor delivers the observed preferences in English. We begin by pursuing the first option. In addition to requiring conjunction, another feature separates the Spanish multi-adjective strings from their English cousins: in Spanish, multi-adjective strings are post-nominal and require conjunction, whereas in English they optionally feature conjunction.
nominal, following the modified noun. Perhaps this feature is responsible for the absence of stable preferences in Spanish. However, Martin (1969) documented stable ordering preferences in Indonesian, a language with post-nominal adjectives. Moreover, the preferences Martin documented matched those found in English, suggesting that Indonesian has subjectivity-based preferences as well. Crucially, while Indonesian has post-nominal adjectives, it does not require conjunction to form its multi-adjective strings. Thus, we find it unlikely that post-nominal adjectives are to blame for the absence of stable preference in Spanish; conjunction appears to be the culprit, and the hierarchical explanations of Simonič and Scontras et al. offer an explanation for why.

But why then should preferences persist in the presence of conjunction in English? Here we are limited to speculation, but there appears to be a viable lead to pursue. Spanish and English differ in that Spanish requires conjunction in multi-adjective strings, whereas in English conjunction is optional. This optionality means that speakers of English will encounter many multi-adjective strings without conjunction, where pressures toward subjectivity-based orderings apply. In other words, the regularity introduced by subjectivity-based preferences is present already in English, such that this regularity can be extended by analogy to cases with conjunction. Here a laundry metaphor seems rather apt: if one washes some white and clean sheets (conjunction) with dirty red socks (no conjunction), the sheets will come out a little pinker after the wash. The regularity present in the conjunction-free cases bleeds over into the cases with conjunction. In Spanish, where multi-adjective strings require conjunction, there is no source for the regularity that could bleed over into the cases with conjunction.

That the English ordering preferences in conjoined multi-adjective strings should find their source in the conjunction-free strings suggests that the preferences do not arise from active, rational reasoning processes from speakers. Rather, these preferences are more likely to arise via an evolutionary process whereby speakers attend to communicative success in the conjunction-free strings, and the regularity that results gets extended by analogy to the conjoined strings. Returning to the predictions we set out to test, our results support the hierarchical explanations of communicative success: with conjunction, pressures for subjectivity-based orderings do not arise so long as the language does not have some other source for ordering regularities.

References


