

## Fake reefs are sometimes reefs and sometimes not, but are always compositional

**Summary.** In semantics, adjective modification is typically handled with set intersection, such that  $[[\text{yellow flower}]] = [[\text{yellow}]] \cap [[\text{flower}]]$ . Thus a *yellow flower* is a *flower*. Such an account, however, runs into problems for adjectives like *fake* or *counterfeit*, which typically have a privative entailment: a *fake fire* is not a *fire* and a *counterfeit dollar* is not a *dollar*. Moreover, privativity cannot easily be encoded as a property of adjectives like *counterfeit*, since e.g. a *counterfeit watch* is judged to be a *watch*, a subsective entailment (Martin, 2022). We gather judgments on over 300 English adjective-noun bigrams (57 novel; i.e., zero corpus frequency), and show that privativity depends on the adjective, noun and context, and can be manipulated for the very same adjective-noun bigram by presenting it in different contexts. This is difficult to explain if privativity is seen as a property of the adjective (del Pinal, 2015; Partee, 2010). Moreover, we find no difference between novel AN bigrams and high frequency ones, suggesting that this is still a case of productive composition and not the result of convention or memorized idiosyncrasy. Our results support compositional accounts like Martin (2022) and Guerrini (2022) which treat privativity as context-dependent.

**Data.** We test 305 adjective-noun bigrams obtained by crossing 38 nouns with 12 adjectives, filtering out bigrams rated to be impossible to assign a meaning in a separate study. 6 typically privative adjectives are matched with 6 typically subsective adjectives of similar corpus frequency: *artificial*, *counterfeit*, *fake*, *false*, *former*, *knockoff*; *homemade*, *illegal*, *multicolored*, *tiny*, *unimportant*, *useful*. The nouns are selected to yield a high quantity of zero-frequency bigrams (19% after filtering), as counted in a  $\sim 200\text{B}$  word corpus (Raffel et al., 2020). Representative high-frequency bigrams include *fake fire* and *counterfeit watch*; zero-frequency bigrams include *fake reef* and *false concert*.

**Experiment 1.** We recruited 510 native English speakers on Prolific (15 excluded). Each participant saw 12 questions (of which 4 fillers) of the form *Is an A N still an N?* (Fig. 1), yielding 10+ ratings/item. Mean bigram ratings are shown in Fig. 3. We find that each “privative” adjective yields graded variation from privative to subsective depending on the noun, and that “subsective” adjectives are less clearly subsective with certain nouns (e.g. *homemade cat*). Further, we find no effect of frequency on rating variance (typ. subsective:  $R^2 = 0.009$ , typ. privative:  $R^2 = 0.014$ ), showing that participants behave similarly for high-frequency and novel adjective-noun bigrams, rather than e.g. having a conventionalized/memorized meaning or entailment only for high-frequency bigrams. Moreover, some zero frequency bigrams like *knockoff image* have quite low variance ( $\mu = 4.90, \sigma^2 = 0.10$ ), showing that participants compose even novel bigrams systematically.

**Experiment 2.** We select 6 pairs of AN bigrams from Experiment 1 with similar middling ratings and high variance, such that one bigram is zero/low frequency and the other is high frequency: *counterfeit diamond/dollar*, *fake reef/fire*, *fake scarf/drug*, *fake glance/plan*, *false concert/war* and *former accusation/house*. For each, we construct two contexts designed to bias the reader towards a subsective or privative entailment respectively (Fig. 2). We recruited 40 native English speakers

Is a counterfeit scarf still a scarf?

Definitely not   Probably not   Unsure   Probably yes   Definitely yes

---

Is a fake gun still a gun?

Definitely not   Probably not   Unsure   Probably yes   Definitely yes

Fig. 1: Sample questions in Exp. 1.

A political party disguises a fundraiser as a concert so that they can hold it at a venue where political rallies aren't allowed. They even hire an up-and-coming band to sing at the event. The false concert is a great success and the attendees enjoy the music as well as networking with the political candidates.

In this setting, is the false concert still a concert?

Definitely not   Probably not   Unsure   Probably yes   Definitely yes

Fig. 2: Subsective-biased context in Exp. 2.

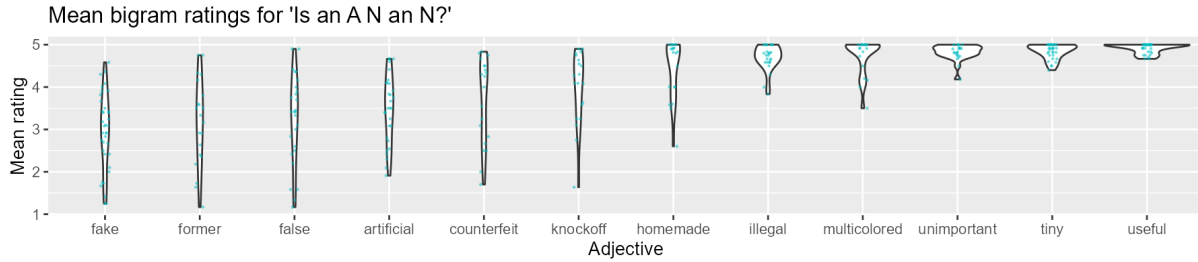


Fig. 3: Mean bigram ratings for Exp. 1, where 1 is most privative and 5 is most subjective.

on Prolific (1 excluded); each participant saw 12 items (of which 6 fillers), yielding 10 ratings/item. We find that for some bigrams (*fake fire*, *fake plan*, *false concert*), the contexts bias participants' entailments very effectively, though other bigrams have more mixed results (Fig. 4) due to item-specific effects (*counterfeit dollar*) or unintended effects of the specific context wording (*fake reef*). We conclude that these entailments are indeed context-dependent and that variation in imagined context may explain some of the variance in Exp. 1. Further, we see no frequency-related patterns in this experiment (e.g. high-frequency bigrams like *fake fire* having less manipulable entailments), showing that deriving entailments from AN bigrams is not conventionalized/memorized and is instead derived from productive use of world knowledge and context. Finally, the ability to manipulate the entailments of novel bigrams such as *false concert* again supports a compositional account.

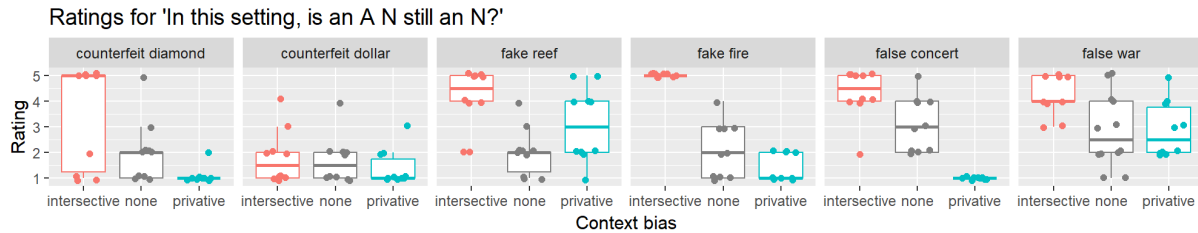


Fig. 4: Selected results for Exp. 2, where 1 is most privative and 5 is most subjective. The ratings from Exp. 1 are shown in gray.

**Discussion.** Our experiments reveal significant variation in privative entailments among so-called privative adjectives and pose problems for any theory (del Pinal, 2015; Partee, 2010) which treats privativity as a property of the adjective. We find that the entailment drawn depends on the adjective and noun (Exp. 1) as well as the context (Exp. 2). This noun and context-dependent variation is equally possible with novel adjective-noun bigrams, and we do not find any effects of frequency/convention, supporting a compositional account of adjective-noun modification nonetheless. One way to capture within-adjective variation without resorting to polysemy is by adapting del Pinal's qualia-based proposal (Martin, 2022): first, all adjectives compose with nouns as functions over noun qualia. For example, *fake* overwrites the telic and agentive qualia of *gun*. A second step evaluates this new bundle of qualia for noun membership to derive subjective/privative entailments. We can adapt this second step to account for context, which influences which qualia matter for determining noun membership. More broadly, the data from these experiments open the door for more detailed accounts which explain how exactly each case of variation is derived.

**References.** • Del Pinal, G. (2015). "Dual Content Semantics, privative adjectives, and dynamic compositionality". *Semantics and Pragmatics* 8 • Guerrini, J. (2022). "Keeping fake simple". *LingBuzz* • Martin, J. (2022). "Compositional Routes to (Non)Intersectivity". PhD thesis • Partee, B. H. (2010). "Privative adjectives: Subjective plus coercion" • Raffel, C. et al. (2020). "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer". *JMLR* 21.140