The effect of standards on scalar implicature processing of gradable adjectives: A web-based eye-tracking study

Properties of measurement scales underlying the meaning of gradable adjectives (e.g., Kennedy, 2007) have been found to affect the availability of pragmatic inferences for these terms (Gotzner et al., 2018). The type of standard value on the measurement scale invoked by gradable adjectives is such a property, which divides gradable adjectives into relative and absolute adjectives: while for relative adjectives the value on the underlying measurement scale that serves as a standard of comparison is contextually determined, for absolute adjectives this is typically a fixed, context-invariant value (Rotstein & Winter, 2004; Kennedy & McNally, 2005). Crucially, it has been argued that scalar implicatures (SIs) of relative adjectives (*warm*$\rightsquigarrow$'warm but not hot') are not derived in all contexts presumably because one needs to be able to resolve the standard for each of the two scale-mates (*warm* vs. *hot*). Absolute adjectives, on the other hand, are more robust SI triggers, especially if the stronger scale-mate is endpoint-denoting (van Tiel et al., 2016), heightening its salience as an alternative for SI computation (Gotzner et al., 2018; Alexandropoulou et al., 2022).

Present study—The present study investigates how the type of standard affects the incremental computation of SIs triggered by gradable adjectives. This will allow us to assess whether lower and upper bounds of gradable adjectives are computed incrementally during compositional interpretation. We build on the visual world (VW) eye-tracking studies by Aparicio et al. (2015, 2018), demonstrating that the processing of relative adjectives hinges on the visual presence of an object (Contrast object) that helps fixing the standard invoked by the relevant adjective (so-called *referential contrast effect* (RCE); cf. Sedivy et al., 1999), whereas the processing of minimum (min) standard absolute adjectives relies solely on linguistic information. Hypothesizing that these semantic differences also factor into the online computation of SIs, we expect to find differential RCEs for relative and min-standard absolute adjectives during incremental interpretation.

Methods—We conducted a web-based eye-tracking experiment using a similar referential communication task to Aparicio et al.'s and the VW paradigm. English native speakers (N=241, recruited from Prolific) were first presented with a visual display of 4 images (see examples in Fig. 1) and 3s later they heard a referring instruction (e.g., *Click on the picture of the warm water with the purple spoon*, Fig. 1(a)/(b)). The instruction is temporarily ambiguous (up to *water* in Fig. 1(a)/(b)) between two referents in the visual scene, i.e., the Target and the Competitor. Importantly, the SI triggered by the adjective in the instruction (*warm*$\rightsquigarrow$'warm but not hot', Fig. 1(a)/(b)) is false of the Competitor, which presents a higher degree of the property encoded by the critical adjective (cf. *warm*). If one were to disambiguate between Target and Competitor by generating the SI associated with the critical adjective of the instruction, this should be reflected in a high(er) proportion of looks to the Target over the Competitor. Participants' task was to click on the correct image after the end of the auditory instruction. Note that the final *with*-PP of the instruction (see Fig. 1) disambiguates the sentence. Participants' eye-movements were collected

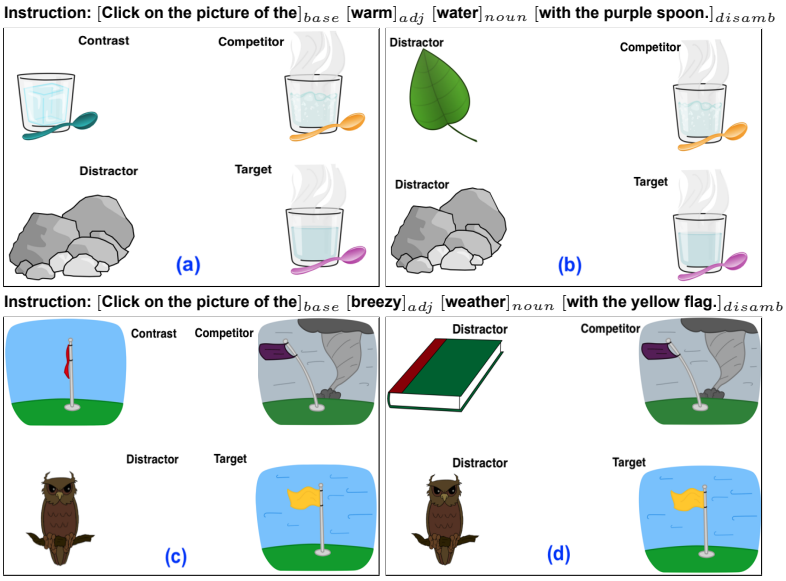

Fig. 1: Example item of relative Horn scale <*warm, hot*> and of min-standard Horn scale <*breezy, windy*> in contrast ((a), (c)) and no-contrast conditions ((b), (d).

from instruction onset until after a selection was made, and were recorded using PCIbex (Zehr & Schwarz, 2018) and the WebGazer.js algorithm (Papoutsaki et al., 2016).

We manipulated the Adjective Type used in the instruction (relative/min-standard) and the presence/absence of a Contrast object in the visual scene (ContrastCond: contrast/no contrast). The Contrast object can be described by the noun (*water*/*weather* in Fig. 1) but not by the adjective of the instruction (*warm*/*breezy* in Fig. 1). We tested 3 relative and 3 min-standard adjective Horn scales (from van Tiel & Pankratz, 2021), the weak scalemate of which has been found by van Tiel & Pankratz to trigger SIs in a picture verification task with pictures like the Competitor images (Fig. 1).

We hypothesize that disambiguation by deriving the SI of the critical adjective of the instruction will be facilitated by processing the comparison standard information of the adjective, and specifically that it will happen differentially for the two adjective types. We predict that disambiguation will be supported by the presence of the Contrast object for relative adjectives, while for min-standard adjectives this should be less likely the case (differential RCE). Therefore, it is expected that participants will fixate on the Target image faster in the contrast condition of relative adjective items than in the respective no-contrast condition, where their looks will be divided between Target and Competitor for longer, whereas such a difference is less likely to be observed between the contrast and no-contrast conditions of min-standard adjectives (Time*AdjectiveType*ContrastCond interaction).

Results—We fit logistic mixed-effects models for three time windows (*adj*(ective), *noun*, *disamb*(iguation)) predicting Target over Competitor looks in terms of time (centered), Adjective Type (sum-coded) and ContrastCond (sum-coded), including the maximal converging random-effect structure justified by our design. Our results revealed a significant 3-way interaction in the *disamb* window (Time*AdjectiveType*ContrastCond: $\beta = 7.62, SE = 2.75, z = 2.78, p < 0.01$), reflecting ongoing processing of ambiguous information. More precisely, this effect reveals that participants converge on the Target faster in the contrast than the no-contrast condition of relative adjectives, while this difference is smaller for min-standard adjectives (see Fig 2).

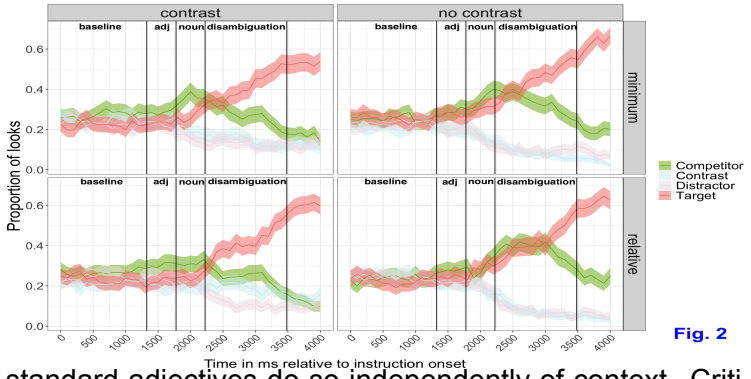

Fig. 2

Discussion—Our finding is in line with our hypothesis: Relative adjectives rely on contextual information to resolve their meaning, while minimum-standard adjectives do so independently of context. Critically, in the contrast condition, the Contrast object lowers the standard for the critical adjective in the relative adjective condition (e.g., *warm*) compared to the no-contrast condition. This happens because in the contrast condition the relevant comparison class includes lower degrees, e.g., of temperature, as compared to the no-contrast condition (see also Barner & Snedeker, 2008; Solt & Gotzner, 2012). Consequently, the degree instantiated by the Competitor is further away from the standard degree for *warm* in the contrast vs. no-contrast condition. In the scalar diversity literature (van Tiel et al., 2016; Gotzner et al., 2018), it is argued that semantic distance is crucial for SI calculation, and a semantically distant alternative to *warm* is highly unlikely to be communicated when uttering *warm*. Hence, when the speaker utters a weak scalar like *warm*, she is more likely to convey that the Competitor degree is excluded in the contrast than in the no-contrast condition.

Overall conclusions—The present study demonstrates that lexical-semantic properties of gradable adjectives are essential to SI processing, and more generally that semantics and pragmatics are highly intertwined during incremental adjective interpretation. We also conclude that web-based eye-tracking may yield fine-grained enough data, advocating for its application in the experimental semantics and pragmatics research.