Description and depiction in a reference game

Most studies of natural language pragmatics focus on felicitous use of **descriptive** modifiers (*red, hungry)*, but we know little about the pragmatics of **depiction**, despite growing acknowledgement of the role of depiction in the compositional semantics of both signed and spoken languages (e.g. Schlenker et al. 2013, Davidson 2015, Henderson 2016, Kuhn & Aristodemo 2017). Here, we investigate the pragmatics of **depictive co-speech gestures**, which have been argued to be compositionally not-at-issue, either as supplements (Ebert & Ebert 2014) or conditional presuppositions/"co-suppositions" (Schlenker 2017), raising many questions about the conditions under which they are felicitous. We focus on the role that triviality/informativity plays in these judgments in the context of a reference game.

Design. Gestures with close equivalents in speech (e.g. *down* in speech and DOWN in gesture) have been the primary focus of previous studies (e.g. Tieu et al. 2017 on gesture projection), but in this project we dissociated *triviality* (duplication of content elsewhere in the proposition) from *informativity* (here, relevance of the content to the reference resolution task) using modifiers that are difficult in English to match across modes: our **verbal modifiers** were descriptive color words (e.g. *blue)*, while our **gestural modifiers** were depictive co-speech gestures of odd shapes (see pictures below). What we varied, through permutations of both distractor items and prompts, was the status of the *informativity* of the modifier, namely whether with a given utterance and stimuli set the target was identified by the modifier in a way that was:

- *Critically informative*: the stimuli pair (e.g. two trolls) that must be differentiated from each other differed in only the dimension (color vs. shape) matching the modifier
- *Informative*: the stimuli pair differed in both dimensions (color and shape)
- *Uninformative*: the stimuli pair differed only in the dimension that matched the unused modifier

In addition, we asked whether these dimensions were judged differently by participants in the verbal descriptive or gestural depictive modes, eliciting both accuracy and felicity judgments.

Data. We recruited 316 self-identified native English speaker participants on Amazon Mechanical Turk, restricted to the United States, whom we directed to Qualtrics questionnaires in two sub-experiments. Each participant viewed 16 trials, each consisting of a short video clip and pair of images of "trolls". For each trial, the participant was asked to i) view a video description of the "target" troll, ii) pick the target troll based on the description (accuracy), and iii) rate the naturalness of the request (felicity rating), using a continuous slider from 0 ("totally awkward") to 1 ("totally natural") (see **Figure 1** for a screenshot).
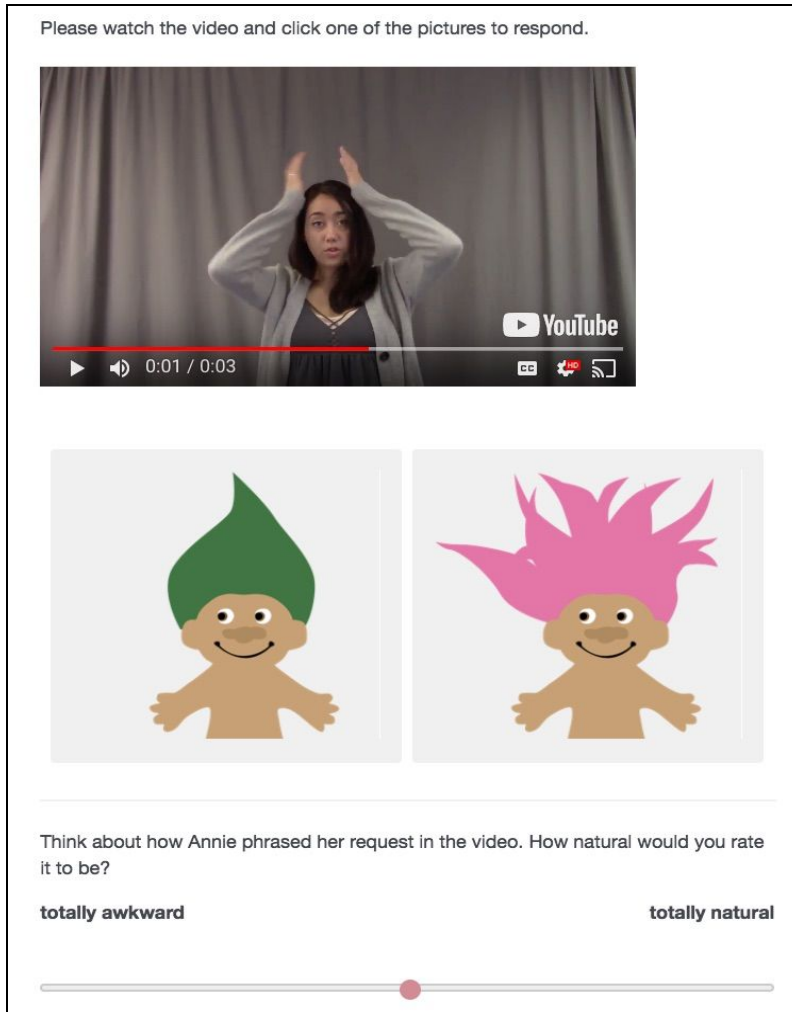
Figure 1. Screenshot of trial from sub-experiment 1.

Descriptions of the target came in four forms, order randomized by participant, varying color and shape, with the shape gestures aligned with *hair*:

a) *Give me the troll with the wild hair.*
   [-linguistic -gesture modifiers]

b) *Give me the troll with the wild* *hair.*
   [-linguistic +gesture modifiers]

c) *Give me the troll with the wild **blue** hair.*
   [+linguistic -gesture modifiers]

d) *Give me the troll with the wild **blue*** *hair.*
   [+linguistic +gesture modifiers]

Wild was included in all trials to reduce both phonological and pragmatic concerns in cases without a linguistic (verbal) modifier (a, b). Trolls varied by *color* (4 total) and *hairstyle* (3 total), which were counterbalanced across four lists across the four cue conditions ((a)-(d) above). Target trolls were always paired with one (counterbalanced) distractor. In sub-experiment 1, participants (N=118) saw distractor trolls that differed from the target troll on

both color and shape, while in sub-experiment 2, participants (N=198) saw distractor trolls that differed either on color, or on shape, but not both; we combine these in Figure 2 below.

Results. Data were analyzed using linear mixed effects models in *R* (*lmer4* package), with participant ID and trial coded as random effects. Independent variables included GestureCue (presence/absence of a gestural modifier), LinguisticCue (presence/absence of a verbal modifier), and Difference (what distinguishes the target and distractor trolls: color/style/both). Felicity rating and accuracy were analyzed as dependent variables.
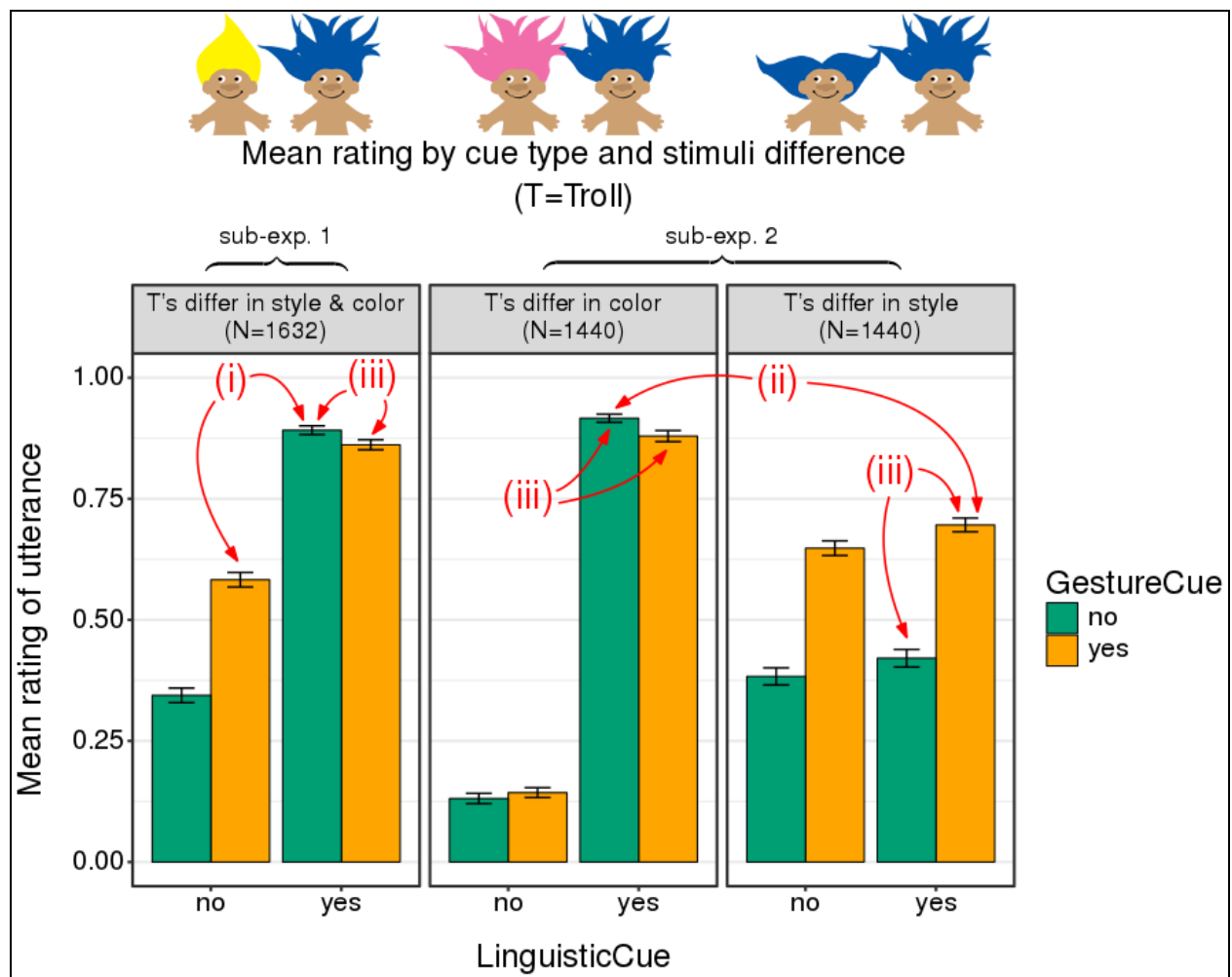

Figure 2. Results comparing the presence (orange bars) and absence (green bars) of gesture cue, based on target/distractor relationship, and presence of linguistic cue.

Summary of Felicity Ratings (see Figure 2)

- **(i)** The gesture-modifier-only condition (b) was rated significantly lower than the linguistic-modifier-only condition (c) when both were informative for reference disambiguation ($p < 0.05$).

- **(ii)** When the linguistic modifier (color word) was critically informative, the highest mean acceptability rating was 0.92, compared to a much lower mean rating of 0.70 when the gestural modifier (hairstyle) was critically informative.
- **(iii)** Overall, adding a gesture to an utterance that contained a linguistic modifier lowered the mean acceptability rating ($p < 0.05$), one exception being in the condition where the gesture was critically informative (i.e. when the troll stimuli differed only in hairstyle).

Summary of Accuracy

- When either the linguistic modifier or the gestural modifier were *critically* informative, participants were significantly more accurate at choosing the target troll ($p < 0.05$) in the former condition (M = 0.98, SD = 0.14) than the latter (M = 0.91, SD = 0.29).

Conclusions. Our findings suggest that both linguistic and gestural modifiers can be used for reference resolution tasks, but may be prioritized differently as linguistic modifiers were both rated as more natural and used more accurately by participants compared to gestural modifiers. The asymmetries we found between gesture and verbal modifiers lead to further questions about the information structural properties of co-speech gestures. The relationship between at-issueness, informativity, and triviality is of particular interest given previous findings based on compositional properties that gestures play a dependent role in discourse, either as supplementary or co-suppositional. Next steps include extending to other communicative contexts and several other types of gesture/verbal pairings, as well as sign languages that can include both color and shape in the same visual mode.

References

Davidson, K. (2015). Quotation, demonstration, and iconicity. *Linguistics and Philosophy 38*(6), 477-520.

Ebert, Co., & Ch. Ebert. (2014). Gestures, demonstratives, and the attributive/referential distinction. Handout of a talk presented at SPE 7, Berlin, June 28, 2014. Available at <http://semanticsarchive.net/Archive/GJjYzkwN>.

Henderson, R. (2016). A demonstration-based account of (pluractional) ideophones. In *Proceedings of SALT 26*, 664-683.

Kuhn, J., & V. Aristodemo. (2017). Pluractionality, iconicity, and scope in French Sign Language. *Semantics and Pragmatics*, *10*(6).

Schlenker, P., J. Lamberton, & M. Santoro. (2013). Iconic variables. *Linguistics and Philosophy 36*(2), 91-149.

Schlenker, P. (2017). Gesture Projection and cosuppositions. *Linguistics and Philosophy* (to appear). Available at <http://ling.auf.net/lingbuzz/002645>.

Tieu, L., Pasternak, R., Schlenker, P., & Chemla, E. (2017). Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa: a journal of general linguistics*, *2*(1):102, 1–27.