Pre-registration report

Study Information

Title: Replication study “Validating the paraphrase methodology” by Scontras & Goodman

Authors: Franziska Scharf, Louis Scheu, Tallulah Jansen, Stefan Warkentin

Description: A problem of ambiguity in semantics and pragmatics is found in plural prediction. This study is particularly concerned with the problem of distributive vs. collective ambiguity (Link, 1983, 1987, 1998; Scha, 1984; Landman, 1989a, 1989b, 1996; Lasersohn, 1988, 1990, 1995, 1998; Schwarzschild, 1994, 1996). E.g., the sentence “the boxes are heavy” can be interpreted distributively, as in “the boxes each are heavy” or collectively, as in “the boxes together are heavy.”
We utilize a forced choice experiment to determine whether certain sentence frames can be used to disambiguate whether a predicate gets pragmatically interpreted in a collective or distributive manner. Using the properties of ambiguous sentence frames and stubbornly distributive predications, we assess whether the words “each” and “together” can be used to unambiguously access distributive and collective semantic interpretations.

Hypotheses:

1. All else being equal, for ambiguous sentence frames (factor-level combination: sentence frame = bare) featuring predicates “big” and “tall”, we expect participants to assign a distributive interpretation reflected in referent choice.
2. All else being equal, for sentence frames featuring "together" (for factor-level combination: sentence frame = together), in utterances such as "The boxes together were tall", we expect participants to assign a collective interpretation reflected in referent choice for the majority of trials.
3. All else being equal, for sentence frames featuring "each" (for factor-level combination: sentence frame = each), in utterances such as "The boxes each were tall", we expect participants to assign a distributive interpretation reflected in referent choice for the majority of trials.

Design Plan

Study type: Experiment

Blinding: Participants are not aware of the experimental conditions. Researchers are not blinded. The present study is administered online, therefore the environment of partaking in the study is identical for each subject.

Study design: We have a mixed 2x3x3 factorial design with factors 1. scenario, 2. predicate and 3. sentence frames. Factors have the following levels: scenario (“move”, “inspect”), predicate (“big”, “heavy, ”tall”), sentence frames (“each”, “together”, “bare”(none)).

Condition 1 is measured between subjects while conditions 2 and 3 are measured within subjects.

After starting the experiment, participants are first presented with either the “move” or the “inspect” variant of the scenario condition. They are informed about the story of an imaginary figure named Pip who has either “moved” or “inspected” boxes in a warehouse. He is either displayed with a cart or without, respectively. Pip is telling another imaginary figure, Jim, about the boxes. Participants are instructed to help Jim understand which boxes Pip refers to during each of the nine trials. Each trial consists of an utterance by Pip such as “*The boxes each were big*”, asking the participant to select (click) one of two distinct images of stacks of boxes, one of which implies a collective interpretation of the utterance (e.g. five small boxes) and one of which implies a distributive interpretation (e.g. two large boxes which were together smaller than the five small boxes). After choosing one set of boxes, participants are immediately forwarded to the next trial. At the end of all nine trials, participants are presented with an optional questionnaire asking “age”, “gender” (with options: male, female, other), level of education (with options: highschool graduate (diploma, Abitur or equivalent), University degree (Bachelor), Higher degree), native Languages: (i.e. the language(s) spoken at home when you were a child) with an open format section and lastly, an option to provide further comments.

If “each” and “together” act as disambiguators for plural prediction, results should indicate an unambious referent that participants choose for trials featuring these words. A distributive “each” should denote the distributive referent while a collective “together” should denote the collective referent. We further expect to see the distributive referent being chosen in ambiguous bare sentence frames featuring the stubbornly distributive predicates “big” and “tall”.

Randomization: Condition 1 is administered in a balanced fashion. Condition 2 and 3 are randomized for all possible combinations on *ad hoc* subject basis amounting to nina trials in total.

Sampling Plan

Existing Data: At submission, a pilot study with a sample of 3 participants has been conducted and analysed. Due to time constraints, data from the main study has already been collected but has not yet been sighted or analysed.

Explanation of existing data: Due to severe time constraints, we were forced to launch the experiment before handing in the preregistration report.

Data collection procedures: We implement the study using the magpie[1] framework and host it on netlify[2].

Requirements for participation are being above the age of 18 at the time of partaking in the experiment and a strong command of the English language. It is explicitly stated that native English speakers are preferred. Participants are recruited via personal contacts (including public whats app and discord groups), mailing lists and the platform reddit. No financial compensation is provided.

Sample size: We aimed to collect a sample of N=30 within the time frame of 72 hours but were able to gather a sample of N=147.

Sample size rationale: We aimed to approximate the sample size N=50 of the original study but reduce it by 20 because of the time constrained recruitment period and lack of compensation. We did not implement a participant limit.

Manipulated variables: In the condition *scenario*, Pip's access to knowledge about the boxes is manipulated between participants by having Pip either "move" or "inspect" the boxes, with or without the display of a cart respectively.

The discourse context is subsequently manipulated within subjects for conditions *predicate* and *sentence frames*. For each trial, Pip utters one of the three predicates "big", "heavy", "tall" in combination with "each", "together", "bare" (none), resulting in 9 utterances in total.

Measured variables: We measure two variables:
1. The chosen arrangement of boxes, whether a collective or distributive meaning was assigned (response/referent choice).
2. The time elapsed between stimuli onset and decision (reaction time) in milliseconds.

Indices: --

Analysis Plan

Statistical models: To analyse our data, we use the lme4[3] package in R. We utilize a mixed effects logistics regression model to predict referent choice by sentence frame("each", "together") and the interaction with predicate(big, heavy, tall) also including trial number. We additionally added the maximal random effect structure warranted by the data: random intercepts for participants and scenarios.

Furthermore, our goal is to evaluate the predicates effect in the sentence frame == bare level taking into account the interaction with the scenario factor.
We thereby again utilize a mixed effects logistic regression model predicting response, based on predicate, scenario and trial number. Additionally, maximal random effects (random intercepts and slopes for participants grouped by trial number) are added.

Transformations: We dummy code the *predicate* predictor with "big" as reference. Depending on the course of our analyses, we may have to code for *sentence frame* and *scenario* as well.

Inference criteria: --

Data exclusion: We exclude data from participants who completed the experiment in less than 300 ms in an effort to eliminate participants who performed accidental clicks or did not read instructions carefully.

Missing data: We only consider subjects if they have completed all tasks. Missing data will be reported as we will report the number of data points used.

---

1 https://magpie-ea.github.io/magpie-site/

2 https://www.netlify.com

3 Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." Journal of Statistical Software, 67(1), 1–48. doi: 10.18637/jss.v067.i01.

Exploratory analysis: We seek to look at the influence of *scenario, predicate* and *sentence frame* on reaction time.

Further, for the ambiguous bare utterances, we will look at referent choice for prompts including “heavy”. We expect a split in collective and distributive interpretation. We also expect *scenario* to influence interpretations of “heavy” since Pip either “moves” or merely “inspects” the boxes. We expect “moves” to yield higher rates of collective referent choice.

Other

Other: --

References

Landman, F. (1989a). Groups I. Linguistics and Philosophy, 12, 559–605.

Landman, F. (1989b). Groups II. Linguistics and Philosophy, 12, 723–744.

Landman, F. (1996). Plurality. In S. Lappin (Ed.), Handbook of contemporary semantics (pp. 425–457). Oxford: Blackwell.

Lasersohn, P. (1988). A semantics for groups and events Ph. D. thesis. The Ohio State University.

Lasersohn, P. (1990). Group action and spatio-temporal proximity. Linguistics and Philosophy, 13, 179–206.

Lasersohn, P. N. (1995). Plurality, conjunction and events. Dordrecht: Kluwer Academic Publishers.

Lasersohn, P. (1998). Generalized distributivity operators. Linguistics and Philosophy, 21, 83–93.

Link, G. (1983). The logical analysis of plurals and mass terms. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), Meaning, use, and interpretation of language (pp. 302–323). Berlin: de Gruyter.

Link, G. (1987). Generalized quantifiers and plurals. In P. Gärdenfors (Ed.), Generalized quantifiers (pp. 151–180). Dordrecht: D. Reidel.

Link, G. (1998). Ten years of research on plurals – where do we stand? In F. Hamm & E. Hinrichs (Eds.), Plurality and quantification. Studies in linguistics and philosophy (Vol. 69, pp. 19–54). Netherlands: Springer.

Scha, R. (1984). Distributive, collective and cumulative quantification. In Truth, interpretation, and information (pp. 131–158). Dordrecht: Foris.

Schwarzschild, R. (1994). Plurals, presuppositions and the sources of distributivity. Natural Language Semantics, 2(3), 201–248.

Schwarzschild, R. (1996). Pluralities. Dordrecht: Kluwer Academic Publishers.