Plural priming revisited: inverse preference and spillover effects

Yizhen Jiang
UCL
yizhen.jiang.19@ucl.ac.uk

Rebecca S. Ren
UCL
rebecca.ren.18@ucl.ac.uk

Yihang Shen
UCL
yihang.shen.19@ucl.ac.uk

Richard Breheny
UCL
r.breheny@ucl.ac.uk

Paul Marty
University of Malta
paul.marty@um.edu.mt

Yasutada Sudo
UCL
y.sudo@ucl.ac.uk

Abstract Maldonado & Chemla & Spector (2017) observe that the distributive interpretation of sentences involving multiple plural expressions gives rise to stronger priming effects than their cumulative interpretation, and propose to interpret this observation in terms of structural priming involving the phonologically silent distributivity operator. We report on a new experiment that included an additional baseline condition, whose results reveal that (i) the observed priming effects are inverse preference effects in that only the less dominant reading in the baseline condition gives rise to sizable priming effects, and (ii) both distributive and cumulative interpretations can have priming effects, depending on speakers' baseline preferences. We argue that these findings undermine Maldonado et al.'s claim that their results offer empirical evidence in support of the existence of a silent distributivity operator in syntax.

Keywords: plurality, distributivity, priming, inverse preference, spillover effects

1 Introduction

Sentences that involve multiple numerical expressions like (1) can receive (at least) two readings, a cumulative reading and a distributive one, as exemplified below.

(1) Two boys have three balloons.
  a. **Cumulative reading**: There are two boys who, between them, have three balloons. Each boy has at least one balloon, and each balloon is owned by at least one boy.
  b. **Distributive reading**: Two boys have three balloons each.

The derivation of a (phrasal) distributive reading like (1b) is standardly assumed to involve a silent *distributivity operator*. For cumulative readings, on the other hand, a number of different compositional semantic theories have been put forward (e.g., Beck & Sauerland 2000; Landman 2000; Schmitt 2019).

Maldonado & Chemla & Spector (2017) conducted a series of priming experiments to investigate the distributive/cumulative contrast. They compared 'distributive primes', which are priming trials forcing the distributive reading, 'cumulative primes', which are priming trials forcing the cumulative reading, and 'neutral primes', which do not force either reading. Their results show that (i) sentences that are ambiguous between a distributive and a cumulative reading were more likely to be interpreted on the distributive reading after distributive primes than after neutral primes (68.33% vs. 61.15%) and (ii) there was no reliable difference in the rate at which the distributive reading was observed after cumulative primes and neutral primes (58.17% vs. 61.15%). Based on these results, the authors make two claims: first, the observed priming effects cannot be characterized as inverse preference effects since the distributive reading was relatively preferred after the neutral primes but still gave rise to priming effects; second, the observed priming effects are to be explained in terms of structural priming of the silent distributivity operator, and as such constitute evidence for its existence.

Maldonado & Chemla & Spector's (2017) experimental design, however, has one potential issue. While the neutral primes were meant to reveal the baseline rate of distributive readings, these primes were interspersed among distributive and cumulative primes and, as previously pointed out for implicature priming (Marty et al. 2022; Waldon & Degen 2020), the presentation of non-neutral primes may trigger strong and long-lasting priming effects which, in turn, may affect the responses in the purportedly neutral priming conditions. If such 'spillover priming effects' were present, then Maldonado & Chemla & Spector's reasoning about the direction of priming could have been misguided.

In the present study, we report on an experiment that is just like Maldonado & Chemla & Spector (2017)'s except that it features additional, more neutral baseline conditions. Specifically, following Marty et al. (2022), we adopted a block design where baseline trials involving no priming at all were conducted in the first block of the experiment and where priming trials were only introduced in the second block. These novel baseline conditions enabled us to obtain data on participants' initial interpretative preferences, i.e., prior to priming. The results indicate that the observed priming effects are actually inverse preference effects, contrary to Maldonado & Chemla & Spector's first claim. Furthermore, we observe that different speakers had different baseline preferences and that, depending on the baseline preferences, both distributive and cumulative readings gave rise to priming effects. We argue that this finding undermines Maldonado & Chemla & Spector's second claim. Finally, comparing our novel baselines and what Maldonado & Chemla & Spector took to be neutral primes, we found evidence for spillover effects.

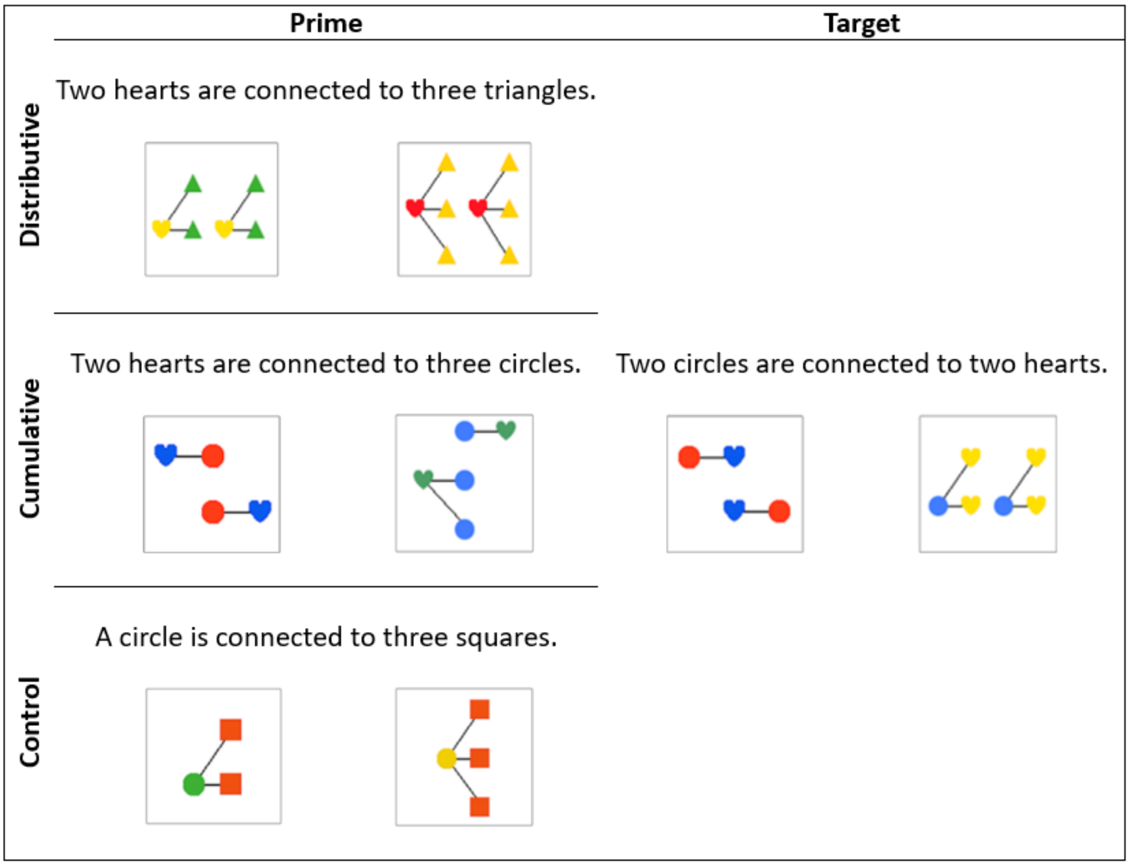
2 Experiment

Our experiment was built upon Maldonado & Chemla & Spector's (2017) materials. It was designed so as to reproduce all the priming conditions from Maldonado & Chemla & Spector's original study while adding novel baseline conditions immune from potential spillover priming effects. To achieve this goal, we adopted a two-block design à la Marty et al. (2022): Block 1 contained no priming trials at all while Block 2 was completely identical to Maldonado & Chemla & Spector's (2017) Experiment 1.

2.1 Participants

75 self-reported native speakers of English (36 females, 38 males, 1 other; average age 40.3 years old) participated in this experiment. Participants were recruited online through Prolific.ac (https://www.prolific.co; see Palan & Schitter (2018) for an overview) using the following pre-screening criteria: English as a first language, UK/US IP addresses, minimum 90% prior approval rating. The survey took around 14 minutes to

Figure 1: Example prime and target trials for the CONTROL, CUMULATIVE and DISTRIBUTIVE conditions tested in Block 2. Participants chose the picture that best fits the sentence. For the prime trials, the expected choice corresponds here to the right picture. For the target trial, if participants interpret the sentence on its distributive reading, they should select here the right picture; otherwise, they should select the left one.



complete and participants were paid £2.4 for their time. All participants gave written informed consent to the processing of their information for the purposes of this study, which was approved by the Research Ethics Committee at UCL. Three participants were excluded prior to data treatment: two for taking the survey on mobile phones rather than computers and one for suspiciously short mean response times (lower than 1s per trial).

2.2 Materials and design

The stimuli were the same as those used in Exp.1 of Maldonado & Chemla & Spector (2017). Each trial consisted of one sentence and two pictures, as exemplified in Figure 1. All sentences were generated based on one of the two frames in (2), where [shape 1] and [shape 2] are placeholders for different shape words from the following list: heart(s), square(s), triangle(s), and circle(s). As reported in Maldonado & Chemla & Spector (2017), the sentence list was generated in Python by randomly inserting shape words and varying the bare numeral quantifier associated with [shape 2] (i.e., *two* or *three*).

(2) a. Two [shape 1] are connected to {two/three} [shape 2]. Ambiguous
b. A [shape 1] is connected to {two/three} [shape 2]. Unambiguous

Block 2 included all the primed target trials and filler trials from Maldonado & Chemla & Spector (2017)'s original experiment. Target trials consisted of an ambiguous sentence obtained from the frame in (2a) and two 'true' pictures, each of which corresponded to one of the readings available for the corresponding sentence (i.e., cumulative or distributive). Each target trial in Block 2 was immediately preceded by two instances of the same prime type: CUMULATIVE, DISTRIBUTIVE or CONTROL. The CUMULATIVE and DISTRIBUTIVE primes all involved an ambiguous sentence obtained from the frame in (2a), as in the target trials. In these primes, one picture was consistent with (exactly) one of the two readings of interest – the cumulative reading in the CUMULATIVE primes and the distributive reading in the DISTRIBUTIVE primes – while the other picture made the sentence false on both readings. The CONTROL primes, previously labelled 'baseline' primes in Maldonado & Chemla & Spector (2017), involved an unambiguous sentence obtained from the frame in (2b). In these primes, one picture made the relevant (unambiguous) sentence clearly false while the other made it clearly false. The filler trials were similar to the prime trials, except that the false picture in these trials used a slightly different visual display. Each of the three priming conditions (i.e., CONTROL, CUMULATIVE, and DISTRIBUTIVE) was instantiated 12 times, giving rise to 48 'Prime-Prime-Target' triplets, and there were 48 fillers, leading to a total of $48 * 3 + 48 = 192$ trials in Block 2.

Recall that our motivation for adopting a two-block design à la Marty et al. (2022) was to establish participants' interpretive preferences prior to the priming phase. For these reasons, Block 1 involves no priming whatsoever: all the trials were either unprimed target trials, corresponding to our BASELINE conditions, or control trials, used as attention checks. These trials were obtained by randomly selecting a third of the target trials and a sixth of the CONTROL primes from Maldonado & Chemla & Spector (2017)'s list of items which we used in Block 2. Thus, Block 1 included 16 unprimed target trials and 16 control trials. We reasoned that, regardless of whether participants initially have a preference for distributive or cumulative readings, the BASELINE conditions from Block 1 should provide us with more pristine baseline rates, against which the experimental conditions from Block 2 can be compared to assess the direction of the priming effects.

2.3 Procedure

The study was run as an online survey using the Gorilla Experiment Builder (https://www.gorilla.sc; see Anwyl-Irvine et al. (2019) for an overview). In the instructions, participants were told that they would be presented with English sentences, each of which would be accompanied by two pictures. They were told that each sentence is meant to describe one and only one of the two pictures and that their task was to decide which picture they think the sentence is describing. They were instructed to provide their responses by clicking on the picture they consider a better match for the sentence.

The instructions were followed by a short practise phase to help the participants get familiar with the visual display and response procedure. During this phase, participants were presented with two practice trials, each of which consisted of an unambiguous sentence and two pictures. They received feedback on their responses and, in case they did not select the right picture, they were asked to redo the trial. Participants could not enter the test phase until they gave correct responses on both practice trials.

The test phase started with the trials from Block 1 (BASELINE conditions) and then continued with the trials from Block 2 (CONTROL, CUMULATIVE and DISTRIBUTIVE priming conditions), with a short self-timed break in-between. Individual and triplet trials were presented in random order in each block. On each trial, a fixation cross appeared and remained on the screen for 500 ms before the items were displayed. For each item, participants provided their response by clicking with the mouse on the picture of their choosing. Items remained on the screen until participants gave their response.

2.4 Data treatment

Responses from one participant were removed from analyses because their performance to the control trials from Block 1 did not reach the pre-established threshold of 80% accuracy. Following much of the recent literature on semantic priming (a.o., Raffray & Pickering 2010; Bott & Chemla 2016; Maldonado & Chemla & Spector 2017; Marty et al. 2022), we further removed all responses to primed target trials that were not preceded by the two correct prime responses. About 2% of the responses to primed target trials were removed through this procedure.

2.5 Data analyses

First, we carried out a global analysis of participants' responses in the target trials to test (a) for pairwise contrasts among the Cumulative, Distributive and Control conditions from Block 2 and (b) for unimodality of the distribution of the by-participant mean rates in the Baseline conditions from Block 1. Participants' baseline rates were found to be distributed bimodally indicating that some participants consistently interpreted the target sentences on their distributive reading while others consistently interpreted these sentences on their cumulative reading. Thus, we carried out a second analysis taking into account participant's baseline preferences: participants were classified as Distributive or Cumulative responders based on their baseline rates, and responses to the target trials were sorted out according to these two responder profiles. Responses from both groups of responders were then analyzed using the data analysis pipelines from the first analysis.

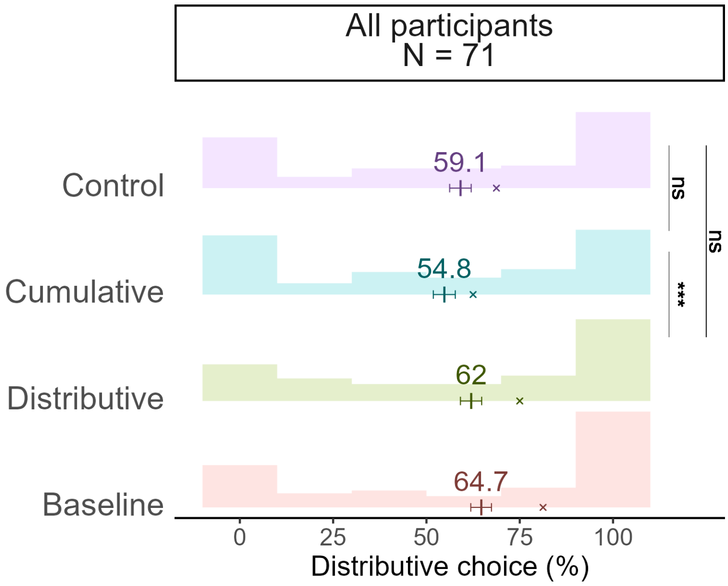
Data analyses were conducted using the `lme4` (Bates et al. 2015), `car` (Fox & Weisberg 2019) and `diptest` (Maechler 2021) libraries for the R statistics program (R Core Team 2021). Participants' response were analysed through model comparison of binomial linear mixed-effects models (Jaeger 2008). The models included random intercepts for Subjects, random slopes for Condition grouped by Subjects and random intercepts for Items (in some cases, only random intercepts for Subjects and Items), as the maximal random effect structure supported by the data. p-values were adjusted using the Bonferroni correction method for multiple testing.

2.6 Results

2.6.1 Global analysis

Figure 2 shows the mean proportion of distributive choices on target trials by experimental condition. Overall, the rates of distributive responses in the priming conditions were similar to those reported in Maldonado & Chemla & Spector (2017), suggesting that participants' performance in the priming trials from Block 2 was largely unaffected by the baseline trials from Block 1. As in Maldonado et al., participants gave significantly more distributive responses in the Distributive than in the Cumulative condition (62% vs. 54.8%; $\chi^2(1) = 8.32$, adjusted-$p = .012$) and no significant difference was found between the Cumulative and Control conditions (54.8% vs. 59.1%; $\chi^2(1) = 4.11$, adjusted-$p = .128$). While the difference between the Distributive and Control conditions did not reach significance ($\chi^2(1) = 2.53$, adjusted-$p = .336$), unlike in Maldonado et al., a trend in the same direction was present, and we consider this small discrepancy between the two studies to be non-essential.

Figure 2: Proportion of distributive choices on target trials by experimental condition. For each condition, the distribution of by-participant mean proportions is visualised by a histogram, the grand mean by a thick bar with its value on top and the 95% CI around it, and the median by a cross. The significance levels are based on the adjusted p-values for all comparisons tested.



Zooming in on the BASELINE conditions, we found that participants' baseline rates were not distributed unimodally ($D = 0.088$, $p < .001$) and that two modes were present in the data: the mode with the highest estimated density value peaked at 98.4% and the second at 4.1%. The estimated location and density values of these modes show that (i) a vast majority of the participants had a strong preference for one of the two readings and (ii) some of them consistently favored the distributive reading while others consistently favored the cumulative reading. In the following, we present a more-fine grained analysis of the data informed by participants' baseline preferences.

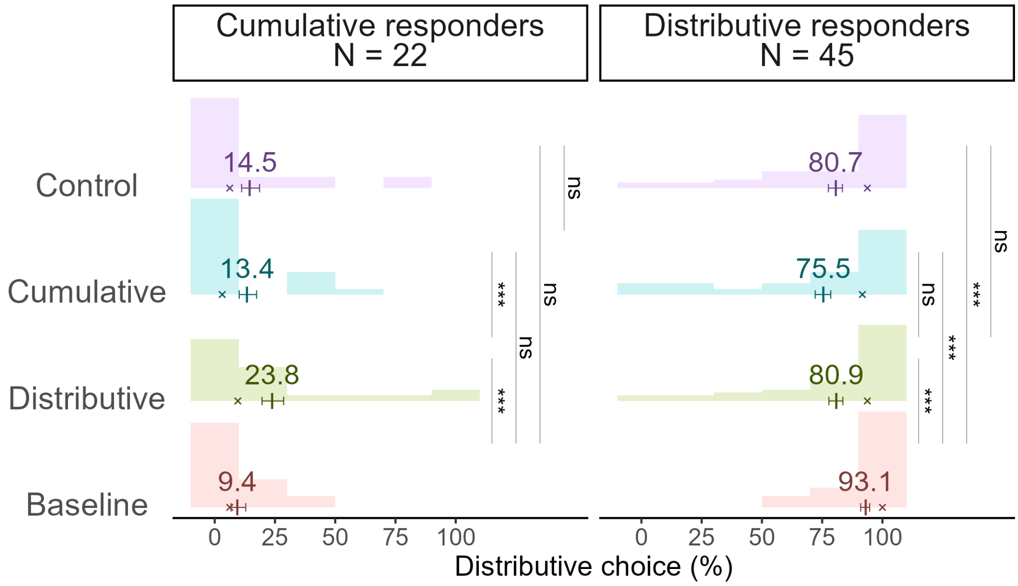
2.6.2 Analysis informed by participants' baseline preferences

In order to see how participants with opposite baseline preferences were affected by the priming conditions, responses to the target trials were sorted according to two responder types: participants were classified as CUMULATIVE responders if their baseline rate was below 50% and as DISTRIBUTIVE responders if their baseline rate was above 50%. There were 22 CUMULATIVE responders and 45 DISTRIBUTIVE responders, representing 30% and 63% of the subjects in our sample, respectively.[1] Figure 2 shows the mean proportion of distributive choices on target trials by responder type and experimental condition.

Pairwise comparisons were conducted between all conditions for each responder type. For the CUMULATIVE responders, the results showed a classical priming effect driven by the DISTRIBUTIVE primes. CUMULATIVE responders gave significantly more distributive responses in the DISTRIBUTIVE conditions (23.8%, 95%CI = [19.6, 28.6]) than in the CU-

[1] 4 participants showed no preference for either reading in the BASELINE conditions. Their results were not included in the following analysis.

Figure 3: Proportion of distributive choices on target trials by experimental condition. This graph reads in an analogous way to the previous one (see Fig.2 for details).



MULATIVE conditions (13.4%, 95%CI = [10.2, 17.4]; $\chi^2(1) = 6.45$, adjusted-p = ...)]. For the DISTRIBUTIVE responders, results showed spillover priming effects from the CUMULATIVE primes onto the other priming conditions. DISTRIBUTIVE responders gave significantly fewer distributive responses in the CUMULATIVE (75.5%, 95%CI = [72.1, 78.5]), CONTROL (80.7%, 95%CI = [77.6, 83.4]) and DISTRIBUTIVE (80.9%, 95%CI = [77.8, 83.6]) conditions than in the BASELINE conditions (93.1%, 95%CI = [91.0, 94.7; all χs> ..., adjusted-ps< ...).

3 Discussion

Our results replicate the findings from Maldonado & Chemla & Spector (2017) in showing that, overall, participants generally provided more distributive responses in the DISTRIBUTIVE than in the CUMULATIVE conditions. But our novel BASELINE conditions also show that there is more to these results than meet the eye. Specifically, we found that both CUMULATIVE and DISTRIBUTIVE primes gave rise to priming effects and that these effects were inverse preference effects, the direction of which depends on speakers' baseline preferences: for participants with an initial preference for the cumulative reading, CUMULATIVE primes had no detectable priming effects whereas DISTRIBUTIVE primes boosted the rate of distributive responses above baseline. Conversely, for participants with an initial preference for the distributive reading, DISTRIBUTIVE primes had no detectable priming effects whereas CUMULATIVE primes decreased the rate of distributive responses across priming conditions. The presence of spillover effects in our data align well with recent work on semantic priming showing that prime trials may have long-lasting effects on participants' decisions (Waldon & Degen 2020; Marty et al. 2022). The detection of such effects in our study also confirms our suspicion that the results from Maldonado et al.'s baseline conditions (our CONTROL conditions) were certainly affected by the presentation of other prime types in the same block of trials.

We argue that the present findings undermine Maldonado & Chemla & Spector (2017)'s claim that their results provide evidence for the silent distributivity operator. First of all, since the results exhibit an inverse preference pattern, we have to conclude that both distributive and cumulative primes have priming effects. If this observation were to be interpreted in terms of structural priming, we would conclude that both readings are to be accounted for by a silent operator, or at least by some distinguished structural property such as silent movement. This is certainly a possibility, but the inverse preference pattern is amenable to an alternative explanation that makes no recourse to structural priming at all. For instance, it can be accounted for in terms of online adaptation of probabilistic expectations about the distributions of the two readings, as proposed for other kinds of syntactic or semantic priming effects (Fine et al. 2010; Marty et al. 2022). Given this alternative explanation, we conclude that, although compatible with the existence of a silent distributivity operator, the results do not provide support for such an operator.

In addition, our study also sheds light on the results of another study, Maldonado & Chemla & Spector (2019), in which the authors used the priming paradigm to investigate the distributive vs. collective interpretations of adjectives in sentences like *The bags are heavy*. The results indicate priming effects from both distributive and collective interpretations, which is, as the authors discuss, unexpected from the view they expounded in Maldonado & Chemla & Spector (2017). However, in this study too, the authors argue that the observed priming effects are unlikely to be inverse preference effects, because the 'baseline' conditions in Experiment 2 indicate an overall preference for the collective interpretation, from which one would expect a larger effect of distributive priming. However, we would like to point out that, as in Maldonado & Chemla & Spector (2017), the baseline conditions used in Maldonado & Chemla & Spector (2019) were interspersed with other priming trials, which could have had spillover priming effects on the baseline trials. Thus, the preference for the collective interpretation could be a byproduct of a relatively stronger priming effect of the collective primes. In addition, they do not report individual variation that might be present in their results, making it difficult to see if there were inverse preference effects in different directions that partially cancelled each other out in the aggregated results. Given these considerations, it is possible that the results they report can be characterised in terms of inverse preference effects, and that any discrepancy between Maldonado & Chemla & Spector (2017) and Maldonado & Chemla & Spector (2019) is attributable to different baseline preferences for the different linguistic stimuli used in these studies. To ascertain whether this is the case, one would have to rerun Maldonado & Chemla & Spector's (2019) experiments with more neutral baseline trials to probe subject's interpretive preferences prior to the priming phase.

In sum, our study replicates the results reported in Maldonado & Chemla & Spector (2017). In addition, our novel baseline conditions uncovered inter-participant variation in the default interpretation, which in turn revealed the inverse preference pattern in the priming results. This means that both distributive and cumulative interpretations can be primed, when they are the dispreferred reading. We pointed out that, as inverse preference effects, the observed priming effects can be explained without referring to structural priming, and therefore do not provide support for the presence of silent operators at LF. The study also provides further suggestions on the priming study of the distributive/collective contrast in Maldonado & Chemla & Spector (2019).

Data availability statement

Stimuli, data, and analysis code are available open access at https://github.com/Yizhen-Jiang/plural_priming_revisited.

Funding information

This research was supported by the Leverhulme Trust Grant RPG-2018-425.

Acknowledgements

We are grateful to Mora Maldonado for sharing with us the original materials used in Maldonado & Chemla & Spector's (2017) experiments.

Competing interests

The author has no competing interests to declare.

References

Anwyl-Irvine, Alexander L. & Massonnié, Jessica & Flitton, Adam & Kirkham, Natasha & Evershed, Jo K. 2019. Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*. 388–407. https://doi.org/10.3758/s13428-019-01237-x.

Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Beck, Sigrid & Sauerland, Uli. 2000. Cumulation is needed: a reply to winter (2000). *Natural Language Semantics* 8(4). 349–371. https://doi.org/10.1023/A:1011240827230.

Bott, Lewis & Chemla, Emmanuel. 2016. Shared and distinct mechanisms in deriving linguistic enrichment. *Journal of Memory and Language* 91. 117–140. https://doi.org/10.1016/j.jml.2016.04.004.

Fine, Alex & Qian, Ting & Jaeger, T. Florian & Jacobs, Robert. 2010. Syntactic adaptation in language comprehension. In *Proceedings of the 2010 workshop on cognitive modeling and computational linguistics*, 18–26. Uppsala, Sweden: Association for Computational Linguistics. https://aclanthology.org/W10-2003.

Fox, John & Weisberg, Sanford. 2019. *An R companion to applied regression*. Third. Thousand Oaks CA: Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Jaeger, T Florian. 2008. Categorical data analysis: away from anovas (transformation or not) and towards logit mixed models. *Journal of memory and language* 59(4). 434–446.

Landman, Fred. 2000. *Events and plurality*. Springer.

Maechler, Martin. 2021. *diptest: Hartigan's Dip Test Statistic for Unimodality - Corrected*. R package version 0.76-0. https://CRAN.R-project.org/package=diptest.

Maldonado, Mora & Chemla, Emmanuel & Spector, Benjamin. 2017. Priming plural ambiguities. *Journal of Memory and Language* 95. 89–101.

Maldonado, Mora & Chemla, Emmanuel & Spector, Benjamin. 2019. Revealing abstract semantic mechanisms through priming: the distributive/collective contrast. *Cognition* 182. 171–176.

Marty, Paul & Romoli, Jacopo & Sudo, Yasutada & Breheny, Richard. 2022. Implicature priming, inverse preference, and context adaptation. Ms., UCL.

Palan, Stefan & Schitter, Christian. 2018. Prolific.ac — a subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17. 22–27.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

Raffray, Claudine N & Pickering, Martin J. 2010. How do people construct logical form during language comprehension? *Psychological Science* 21(8). 1090–1097.

Schmitt, Viola. 2019. Pluralities across categories and plural projection. *Semantics and Pragmatics* 12. 17.

Waldon, Brandon & Degen, Judith. 2020. Symmetric alternatives and semantic uncertainty modulate scalar inference. In Denison, Stephanie & Mack, Michael & Xu, Yang & Armstrong, Blair C. (eds.), *Proceedings of the 42nd annual meeting of the cognitive science society*.