Plural priming revisited: *inverse preference* and *spillover effects*[*]

Yizhen Jiang[1], Rebecca S. Ren[1], Yihang Shen[1], Richard Breheny[1], Paul Marty[2], and Yasutada Sudo[1]

[1] University College London, London, UK
[2] University of Malta, Msida, Malta

Abstract

Maldonado et al. (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 of the phonologically silent distributivity operator. We report on a new experiment that included an additional 'neutral' 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 evidence the existence of the silent distributivity operator in syntax.

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 (1-b) 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., [3, 11, 15]).

Psycholinguistic studies have raised evidence that distributive readings are generally less preferred and involve more processing effort in comparison to cumulative and other non-distributive readings [1, 4, 5, 7, 10, 16, 17]. Particularly relevant for the present study is [12], which used priming. Their experimental results show (a) that sentences that are ambiguous between distributive and cumulative readings are more likely to be interpreted on the distributive reading after priming trials that force it (distributive primes) in comparison to the unprimed target trials that involve no priming (68.33% vs. 61.15%); and (b) that after priming trials that force the cumulative reading (cumulative primes), the rate at which the distributive reading is observed is not significantly different from the unprimed trials (58.17% vs. 61.15%). Based on these results, they make two claims: (i) The observed priming effects cannot be characterized

[*]We are grateful to Mora Maldonado for sharing with us the materials used in Maldonado et al.'s (2017) experiments. Data and analysis code are available at https://github.com/Yizhen-Jiang/plural_priming_revisited. This research was supported by the Leverhulme Trust Grant RPG-2018-425.

as inverse preference effects since the distributive reading was relatively preferred in the unprimed condition (61.15%); and (ii) 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.

[12]'s experimental design, however, has one potential issue: They obtained their baseline results from trials that were interspersed between trials that involved priming, but as previously pointed out for implicature priming [13, 18], such baseline trials are potentially affected by preceding priming trials and do not directly inform us about neutral preferences.

In the present study, we report on an experiment that is just like [12]'s except that it has more neutral baseline conditions. Specifically, following [13, 18], we adopted a block design where the baseline trials with no priming were all conducted in the first block of the experiment, and priming was only introduced in the second block. These novel baseline conditions enabled us to obtain data on participants' natural interpretative preferences. The results indicate that the observed priming effects are actually inverse preference effects, contrary to [12]'s claim (i) mentioned above. Furthermore, we observe that different speakers had different baseline preferences, and depending on the baseline preferences, both distributive and cumulative readings gave rise to priming effects. We argue that this finding undermines [12]'s claim (ii).

2 Experiment

Our experiment was built upon [12]'s materials and design while adding a more neutral baseline condition in order to investigate the direction of the priming effects observed in [12], which [12] could only speculate about in the absence of proper baseline. To achieve this, we adopted a two-block design: Block 1 contained no priming trials at all, allowing us to establish participants' interpretative preferences without potential spillover effects; Block 2 contained the same trials as in [12]'s Experiment 1. With this two-block design, the baseline conditions are protected from potential contamination effects due to spillover from prime trials, as suggested in [13].

2.1 Participants

75 self-reported native speakers of English (36 females, 38 males, 1 other; average age 40.3 years old) were recruited via Prolific (UK/US IP addresses; prior approval rate $\geq$ 90%). Three participants were excluded prior to analyses: two for taking the survey on mobile phones rather than computers and one for suspiciously short mean response times (lower than 1s per trial). The survey took around 14 minutes to complete and participants were paid £2.4 for their time. All participants gave written informed consent.

2.2 Materials and Design

Our experimental materials were identical to Experiment 1 of [12]. Each trial consisted of one sentence and two pictures as exemplified in Figure 1. All sentences in the trials were generated based on one of the two frames as (2). Terms used for [shape 1] and [shape 2] featured the following shape names: heart(s), square(s), triangle(s), and circle(s). In each sentence, the name for [shape 1] was always different from that of [shape 2]. As reported in [12], the sentence list was generated in Python by randomly inserting shape words. To illustrate, Figure 1 presents schematized primes and targets used in the experiment.

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

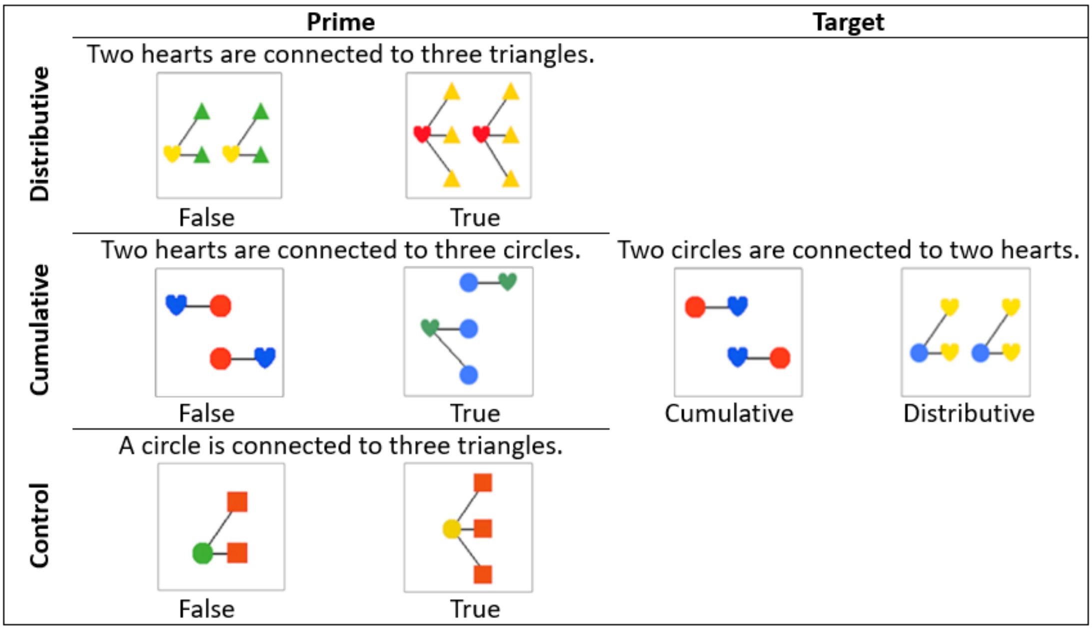


Figure 1: Example prime and target trials in the Cumulative, Distributive, and Control conditions in Block 2

b. A [shape 1] is connected to {two/three} [shape 2]. Unambiguous

Prime trials in Block 2 consisted of an ambiguous sentence and two pictures: one picture was consistent with one of the two readings (cumulative readings for cumulative primes and distributive readings for distributive primes) while the other was false because of the wrong number of shapes regarding of sentence. Control trials (called 'baseline' trials in [12]) consisted of an unambiguous sentence together with two pictures: one picture made the sentence true while the other made it false. As for target trials, sentences were all ambiguous. In the meantime, the two pictures were both "true" as they each corresponded to one of the readings of this sentence. Filler trials were similar to prime trials, except the false picture employed a different visual display.

Block 2 intended to replicate the original experiment in [12]. trials in this block followed a triplet pattern, meaning there were always two cumulative primes, two distributive primes, or two control primes before each target trial. There were 48 triplets, meaning 144 experimental trials in total. Another 48 filler trials were added, which made the total number of trials in this block achieve 192.

Trials in Block 1 were extracted from trials in block 2. Target trials in this block were the same as target trials in Block 2, which consisted of one ambiguous sentence and two pictures (one was consistent with cumulative reading, and another was consistent with distributive reading). At the same time, filler trials were the same as control trials of Block 2.

Block 1 included 16 target trials, and 16 filler trials were inserted randomly. In this block, we intended to establish a neutral baseline without any possible priming effects. Participants could choose the picture according to their interpretations of sentences. In this block, different participants might exhibit different preferences between the cumulative reading and distributive reading. Whether they were pro-cumulative or pro-distributive people, their statistics would serve as a better baseline for assessing priming effects' direction.

2.3 Procedure

Participants were instructed to click on the picture which they consider to be a better match for the sentence. Then followed a short practice phase where participants were given two practice trials, each of which contained an unambiguous sentence accompanied by two pictures. Feedback was given on participants' responses. Additionally, participants were asked to redo the trial if they select incorrect images. Participants could not enter the test phase until they gave correct responses on both practice trials. After the practice phase, the experiment began with the trials from Block 1 (Baseline conditions) and then continued with the trials in Block 2 (Distributive, Cumulative, and Control conditions).

2.4 Data treatment

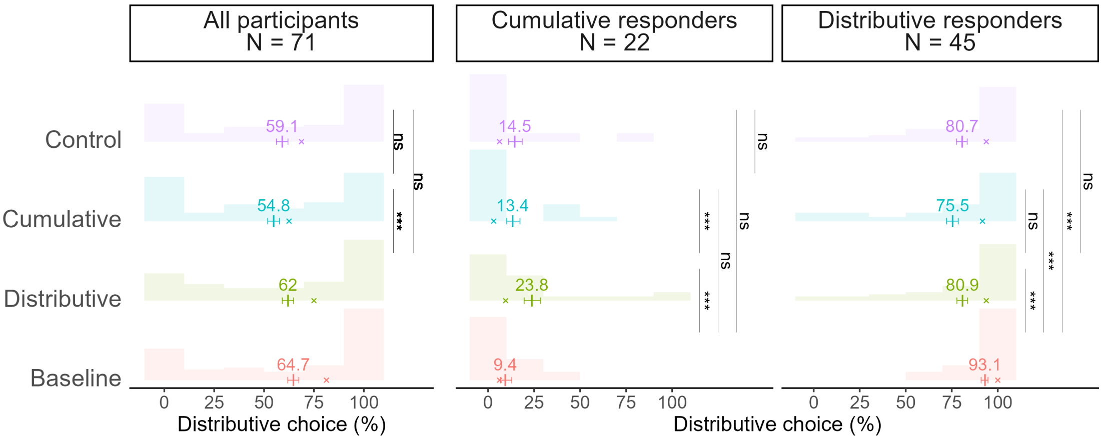
Responses from one participant were removed from analyses because their mean accuracy rate of primes and fillers was below the pre-established 80% threshold. Following [12, 14], we further removed all responses to primed Target trials that were not preceded by the two correct prime responses. About 2% of all target responses were removed due to incorrect prime responses.

2.5 Results

The proportion of distributive choices on target trials by condition for all participants is displayed in Figure 2.5a. We conducted pairwise comparisons [9] between Distributive, Cumulative, and Control conditions in Block 2. Target responses from each group were analyzed by modelling response type likelihood using logit mixed effects models with the maximal random effects structure [2] which included random intercepts for subjects and items. The results of Block 2 showed essentially the same pattern as in [12]: Participants made significantly more distributive responses in the Distributive condition than in the Cumulative condition (62.0% vs. 54.8%; adjusted-$p < .001$) and no significant difference was detected between the Cumulative and Control conditions (54.8% vs. 59.1%; adjusted-$p = 0.154$). While the difference between the Distributive and Control conditions did not reach significance, unlike in Maldonado et al., the tendency in the same direction is present, and we consider this discrepancy between the two studies to be non-essential. This suggests participants' performance in the Priming block was largely unaffected by the prior Baseline block and hence secures our assessment of the participants' behavior in the Priming conditions.

Zooming in on the Baseline block, our analysis of the Baseline results using Hartigan's dip test of unimodality [8] identified two-mode distribution in participants' baseline preferences. This reveals that the majority of the participants had a preference for one of the two readings: Some participants (n = 22) tended to interpret the target sentences on their cumulative readings (henceforth, cumulative responders) while others (n = 45) favored the distributive interpretation (henceforth, distributive responders). The rest (n = 4) showed no preference.

In light of the inter-speaker baseline variation, we conducted further analysis of the results of Block 2 by sorting them according to the responder type (see Figure 2.5b). Pairwise comparisons [9] between Baseline, Distributive, Cumulative, and Control conditions were conducted for each responder group. Similar to the global analysis, target responses from each group were analyzed by modeling response type likelihood using logit mixed effects models [9] with random intercepts for subjects and items. The results reveal that the priming effects are **inverse**



(a) (b)

Figure 2: Proportion of distributive choices on target trials (a) by condition for all participants and (b) by responder group and condition. Histograms show the distributions of by-participant mean proportions. Crosses denote medians. Thick lines with numerals on top denote grand means. Error bars denote 95% binomial proportion confidence intervals. Asterisks denote significance levels based on the adjusted-p values for all comparisons.

preference effects in the following sense: For the cumulative responders, the priming effects were driven by the distributive primes boosting the rate of the distributive interpretation, while for the distributive responders, the priming effects were mainly driven by the cumulative primes inhibiting the rate of the distributive interpretation.

Thirdly, we observe spillover effects. Specifically, the distributive responders gave significantly fewer distributive responses in the Control condition than in the Baseline condition (80.7% vs. 93.1%; adjusted-$p < .001$), which confirms our suspicion that Maldonado et al.'s "baseline" results were not neutral, and affected by spillover effects from other trials in the same block. Furthermore, a similar difference was observed between the Distributive condition and the Baseline condition for the distributive responders (80.9% vs. 93.1%; adjusted-$p < .001$). This is consistent with the inverse preference pattern, where the distributive primes, forcing the favored interpretation for this type of participant, had no priming effect, and the difference from the Baseline condition was due to spillover effects. For the cumulative responders, we have a general tendency towards the opposite pattern, but the differences did not reach significance. We suspect this is partly due to the fact that there were relatively few cumulative responders in our data set.

Further, we detect no priming effects of the preferred interpretation on its own. Our results showed that the rate of distributive choice was about the same between the Control and Distributive condition for distributive responders and also between the Control and Cumulative condition for cumulative responders. It suggested that the effects observed in the Cumulative condition for cumulative responders were all attributed to the spillover effects from the distributive primes. Hence, cumulative primes could hardly modulate cumulative responders' interpretation. Similarly, distributive primes have no priming effects for distributive responders. That is, the favored interpretation in the Baseline condition did not have any observable priming effects by itself.

Finally, note that the size of the spillover effects is always smaller than the size of the priming effects. This is because spillover effects are in fact long-lasting remaining priming effects. Therefore, the difference between the Distributive and Cumulative conditions observed for all participants (interpreted as priming effects in [12] was the sum of the difference between the spillover effects and priming effects of Distributive primes for cumulative responders and that of Cumulative primes for distributive responders.

3 Discussion

Our results replicate the priming effects observed in [12] that participants provided more distributive responses after distributive primes than after cumulative primes. In addition, we found that these global effects are in fact caused by different prime types, depending on participants' baseline interpretive preferences. Specifically, distributive primes boosted the rate of distributive responses for cumulative responders while cumulative primes boosted the cumulative interpretation for distributive responders. We further observed that robust priming effects spilled over the targets in the other conditions within the same block: In both Control and Distributive conditions, cumulative responses were boosted by the lasting priming effects from cumulative primes in the Cumulative condition. Finally, we did not detect any priming effects of the less preferred interpretation: no additional effects were found in the Distributive condition relative to the Cumulative condition for distributive responders.

We argue that the above observations undermine [12]'s claim about the distributivity operator. Firstly, 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 some distinguished structural property such as silent movement), not just the distributive reading. Secondly, the inverse preference pattern is amenable to an alternative explanation without recourse to structural priming. Specifically, it could be accounted for in terms of online adaptation of probabilistic expectations about the distributions of the two readings, as proposed for certain other priming phenomena [6, 13]. 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.

Our study replicates the results reported in [12]. In addition, our more neutral Baseline condition uncovered inter-participant variation in the default interpretation, which in turn revealed the inverse preference pattern in the priming results as well as spillover effects. This means that both distributive and cumulative interpretations can have priming effects, when they are the dispreferred reading. We pointed out that the observed priming effects as inverse preference effects can be explained without referring to structural priming, and therefore do not provide support for the presence of silent operators at LF.

References

[1] Stavroula Alexandropoulou, Jakub Dotlačil, Yaron Mcnabb, and Rick Nouwen. Pragmatic inferences with numeral modifiers: Novel experimental data. In *Proceedings of SALT 25*, pages 533–549, 2015.

[2] Dale J Barr, Roger Levy, Christoph Scheepers, and Harry J Tily. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3):255–278, 2013.

[3] Sigrid Beck and Uli Sauerland. Cumulation is needed: A reply to Winter (2000). *Natural Language Semantics*, 8(4):349–371, 2000.

[4] Patricia J. Brooks and Martin D. S. Braine. What do children know about the universal quantifiers all and each. *Cognition*, 60:235–268, 1996.

[5] Jakub Dotlacil. *Anaphora and distributivity: A study of same, different, reciprocals and others.* Netherlands Graduate School of Linguistics, 2010.

[6] Alex Fine, Ting Qian, T. Florian Jaeger, and Robert Jacobs. Syntactic adaptation in language comprehension. In *Proceedings of the 2010 Workshop on Cognitive Modeling and Computational Linguistics*, page 18–26, Uppsala, Sweden, Jul 2010. Association for Computational Linguistics.

[7] Lyn Frazier, Jeremy M Pacht, and Keith Rayner. Taking on semantic commitments, II: collective versus distributive readings. *Cognition*, 70:87–104, 1999.

[8] John A Hartigan and Pamela M Hartigan. The dip test of unimodality. *Annals of Statistics*, pages 70–84, 1985.

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

[10] Barbara Kaup, Stephanie Kelter, and Christopher Habel. Representing referents of plural expressions and resolving plural anaphors. *Language and Cognitive Processes*, 17(4):405–450, 2002.

[11] Fred Landman. *Events and Plurality.* Springer, 2000.

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

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

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

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

[16] Kristen Syrett and Julien Musolino Rutgers. Collectivity, distributivity, and the interpretation of plural numerical expressions in child and adult language. *Language Acquisition*, 20(4):259–291, 2013.

[17] Cherlon Ussery. Processing plural DPs: Collective, cumulative, and distributive interpretations. In *UMOP*, volume 37. GLSA, 2007.

[18] Brandon Waldon and Judith Degen. Symmetric alternatives and semantic uncertainty modulate scalar inference. In Stephanie Denison, Michael Mack, Yang Xu, and Blair C. Armstrong, editors, *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*, 2020.