Supplemental Materials for *The logic in language: How all quantifiers are alike, but each quantifier is different*

Context effects: Effects of Prime Quantifier independent of Prime Scope

In Experiment 2, we reported the presence of within-quantifier priming and the absence of between-quantifier priming. In addition to the priming effects -- effects of a prime trial on the target that follows it -- seeing sentences with the same prime quantifier repeated throughout the experiment may also have a contextual effect on target choice, independent of the prime scope in a specific prime-target pair. As Experiment 1 shows, some quantifiers are strongly biased toward a U-wide interpretation (EACH being the most biased) and others are heavily E-wide biased (the numbers most so). Regardless of participants' preferred reading, half of the prime trials a participant sees always force a U-wide reading and the other half force an E-wide reading. This gradually gives participants increasing evidence that the normally dispreferred reading may be acceptable for that quantifier, and possibly for other quantifiers that are represented as similar. Thus, if participants repeatedly see a quantifier that they would normally assign either a U-wide or E-wide interpretation in the context of prime trials where the only available reading is the dispreferred one, they may adjust their expectations over time, resulting in a weaker bias on subsequent target trials containing that quantifier. Depending on the relevant dimension of similarity between quantifiers, this effect may extend across some or all quantifiers in the Between-Quantifier conditions of Experiment 2. In all cases, we would expect such a context effect to push target trial responses away from the baseline for that individual quantifier and closer to chance.

The key question is whether such adaptation effects exist in the present experiment, and if they do, whether speakers adapt their expectations about the use of one quantifier based only on distributional information about that same quantifier, or whether some pairs of quantifiers share a representational similarity that is relevant for aggregating probabilities across them. In other words, although we do not find between-quantifier *priming* effects, perhaps there are systematic between-quantifier *context* effects that reveal some representational overlap between certain quantifiers.

1.1.Results

1.1.1. Within-Quantifier Context Effects

In the Within-Quantifier conditions, participants see the same quantifier word on subsequent trials – within each prime-target pair – as well as repeated throughout the experiment. Since each quantifier has a strong baseline bias, experience with primes that only match this bias half of the time could shift performance toward chance. Figure S.1 shows the Within-Quantifier conditions in Experiment 2 relative to their corresponding baselines in Experiment 1. For each target quantifier, we constructed a separate model comparing the baseline data from Experiment 1 to the Within-Quantifier priming condition for that word in Experiment 2. The models included a between-subjects two-level variable of Prime Presence as the main predictor of interest (for example, the EVERY baseline from Experiment 1 compared to the EVERY-to-EVERY priming condition in Experiment 2).

There was a significant effect of Prime Presence for every quantifier, except in the ALL-to-ALL condition (z=1.61, p=0.1), suggesting that participants were sensitive to the presence of primes where they were forced to pick a dispreferred interpretation. We replicated the ALL-to-ALL Within-Quantifier condition with an additional 256 participants, of which 235 were included, after applying the usual exclusion criteria. Including both the original run and replication data, we first fit a logistic mixed-effects model to look for the effect of Prime Type in this condition, but adding a two-level fixed effect of replication (Replication vs. Original Run) and found no effect of replication, nor any interaction between prime type and replication. In this data, we replicated a highly significant effect of prime type (U-wide or E-wide), with $p<0.0001$ (see Sec 3.2.1.). Pooling both the replication and original data, we fit a model comparing the Within-Quantifier ALL condition to the ALL baseline from Experiment 1. Including the additional replication data, we now found a significant effect of Prime Presence ($z=2.51$, $p=0.01$), indicating a significant difference between the Within-Quantifier priming ALL-to-ALL condition and the ALL baseline condition. Consistent with the results for the other quantifiers, the priming condition was closer to chance (23% mean U-wide responses) than the baseline condition (21% U-wide responses). Given this pattern of results, we conclude that the original failure to find an effect in the ALL-to-ALL condition was likely a false negative. In sum, we have some evidence of context effects in all of the Within-Quantifier priming conditions, pushing the average bias closer to 50/50.

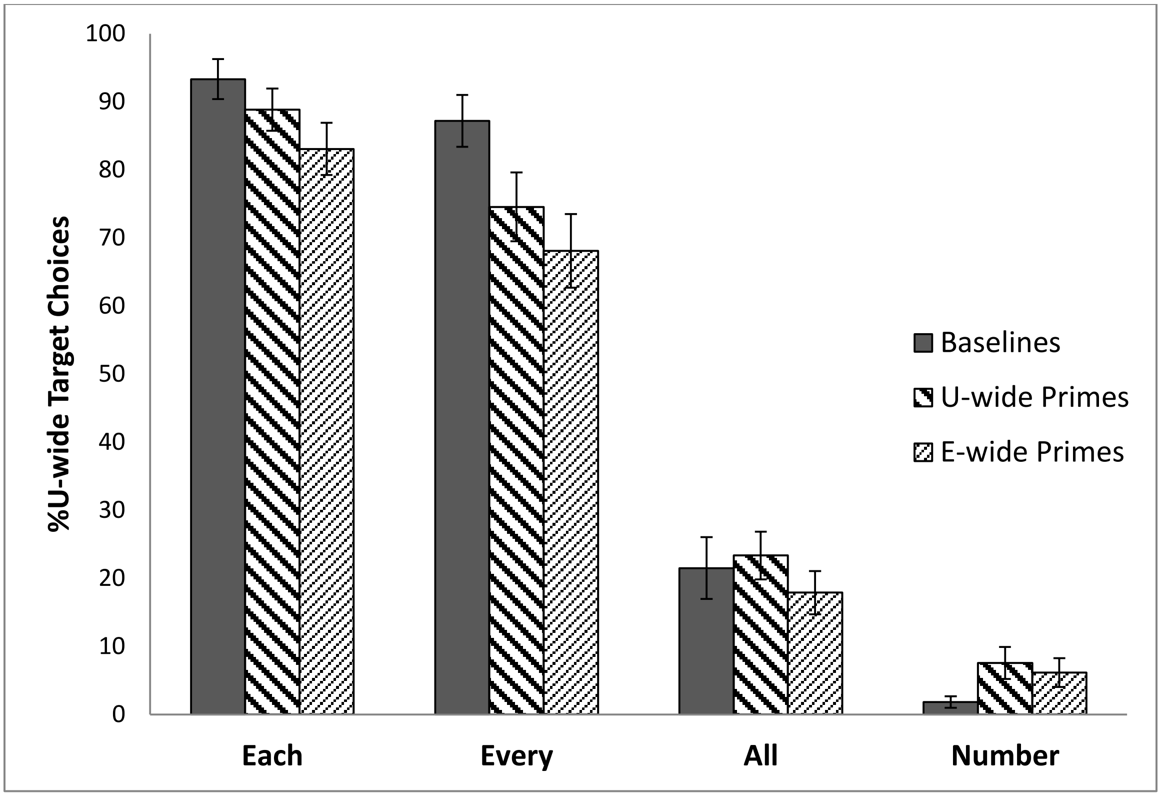

Fig. S.1: The Within-Quantifier conditions of Experiment 2 alongside the baselines from Experiment 1. Error bars indicate 95% confidence intervals, averaged across items with subjects as the random variable. The U-wide and E-wide prime bars in the ALL-to-ALL condition do not include data from the replication.

1.1.2. Between-Quantifier Context Effects

Given that distributional information about a quantifier's scope preference is accumulated and stored, we can now ask which sets of quantifiers this information is aggregated over. Does information that changes the bias of one quantifier change the bias of another? It could be that probabilities get updated only for a specific quantifier word, in which case we would expect to see main effects of Prime Quantifier in the Within-Quantifier conditions, but not in Between-Quantifier conditions, just as with the priming effects. Alternatively, it could be that two or more quantifiers form a single class, such that evidence that one quantifier is more or less biased than expected would shift the bias of another quantifier in the same class. Finally it is possible that these probabilities are aggregated across all of the quantifiers as a single class, in which case there should be effects of Prime Quantifier in all of the Between-Quantifier conditions.

We report and discuss the absence of Between-Quantifier priming effects in Experiment 2 in section 3.2.2 of the paper. To look for these priming effects, we constructed four separate models, one for each target quantifier, each of which included the following predictors: a three-level fixed effect of Prime Quantifier (capturing the three different Between Quantifier conditions), a two-level fixed effect of Prime Scope, and their interaction. When looking for priming effects, we were interested in the interaction between Prime Scope and Prime Quantifier and in the main effect of Prime Scope. However, to look at the effect of prime sentences on the subsequent targets independent of the type of prime, we are interested in the main effect of Prime Quantifier. To look at these main effects, we compare the fit of a model including only the fixed effect of Prime Scope to a model including Prime Scope and Prime Quantifier (with no interaction term), repeating this for each target quantifier. Across these four model comparisons, we find two comparisons where the effect of Prime Quantifier is significant: across the conditions where EVERY is the target quantifier ($\chi^2(2)=11.6$, $p=0.003$) and those where EACH is the target ($\chi^2(2)=16.83$, $p=0.0002$). Looking at the fixed effects in this model using a dummy coding scheme where the ALL- is treated as the reference level, we find that in the model where EVERY is the target, there is a main effect of the condition where EACH is the prime (z=-2.91, $p=0.004$), and in the model where EACH is the target, there is a main effect of the condition where EVERY is the prime (z=-3.12, $p=0.002$). To follow up on these effects, we compared these conditions to their respective baselines for those target quantifiers (from Experiment 1). In a model including just the EVERY baseline condition from Experiment 1 and the EACH-to-EVERY priming condition from Experiment 2, we find a highly significant effect of Prime Presence (z=-

5.99; p<0.0001). Similarly, in a model including the EACH baseline condition from Experiment 1 and the EVERY-to-EACH priming condition from Experiment 2, we find a significant effect of Prime Presence (z=-3.02; p=0.003). These effects are parallel to the Within-Quantifier context effects: the presence of a prime pushes target responses down closer to 50-50 and away from the quantifier's bias.

The experiments we report were designed to look at priming effects using a within-subject manipulation of the prime type. Aside from the different between-subjects conditions in Experiment 1, which were run simultaneously with random assignment of participants, no effort was made to ensure random assignment in the other experiments. Therefore, the context effects we found here depend on a between-subjects comparison of conditions without random assignment. Furthermore, the effects seem less robust than the priming effects, which we consistently replicate and report in the main paper. The effect of Prime Quantifier in comparing the ALL-to-ALL condition to the ALL baseline was not found in the original dataset, but only upon replication. While these effects suggest the possibility of speakers adapting their expectations about a quantifier's baseline bias based on information from the scope readings available in other sentence, to directly establish the presence of speaker adaptation effects would require a different design that would address this question directly.

Only tentatively, then, we interpret these effects as evidence that participants adjust their expectations about the scope bias of EACH based on distributional information about EVERY, and vice versa. Why would this information get aggregated across this pair of quantifiers and not across other pairs? Perhaps there is some semantic property that is shared by EVERY and EACH, making them more similar to one another than they are to ALL or to numbers (see Champollion, 2010; Steedman, 2012, and Sec. 6.3. and 6.4. for theoretical considerations along these lines).

Note, however, that any similarity of this kind is apparently not sufficient to allow *scopal* priming between EVERY and EACH. The systematic lack of Between-Quantifier priming effects between EACH and EVERY and the presence of Between-Quantifier context effects for this pair suggests that there are at least two distinct types of representations or processes involved in LF construction (operating over different time scales), which our theoretical accounts should address. Constructing an LF online is a process that draws on prior beliefs about the scope bias of the quantifiers involved. The process that constructs the representation of quantifier scope (presumably based on the instructions provided by these priors) seems to treat these quantifiers as wholly distinct. In contrast, the process that updates these prior beliefs appears to be sensitive to a representational similarity between EACH and EVERY.

Discussion

Separate from priming, we find a context effect that occurs when the quantifiers in the primes and the targets are the same, and between the quantifiers EACH and EVERY. In these cases, repeated experience with primes that illustrate both scope readings shifts our participants' interpretation of the target sentences in the direction of chance (relative to the baseline) suggesting that their expectations for the scope of target quantifier has changed.

Participants' shifting expectations about a quantifier's likely position at LF can be understood in the context of other semantic adaptation effects. When exposed to a speaker using SOME and MANY to refer to proportions of sets differently than how a listener otherwise would, listeners change their expectation about what a speaker is likely to mean by these words in the future (Yildrim, Degen, Tanenhaus, & Jaeger, 2015). Such adaptation effects have been found not only for the meanings of individual words but also for ambiguous syntactic structures, with listeners resolving a speaker's ambiguous utterance consistently with that speaker's past usage of syntactically similar utterances in less ambiguous contexts (Kamide, 2012). The present work is the first to show such an adaptation effect for a semantically ambiguous structure at the level of LF.

The direction and size of the context effects we find in the present data are also consistent with this literature. All of the context effects we find shift a quantifier's bias towards the 50-50 mark, indicating a shift towards the less common interpretation and suggesting that the prime trials that force that interpretation have a stronger effect than those trials that force an interpretation consistent with the preexisting bias. The least biased quantifier, ALL, showed the smallest and least stable within-quantifier context effect. This suggests that listeners learn a speaker's quantifier bias more from generally unexpected utterances than from expected ones, consistent with adaptation effect found in syntactic priming (Fine & Jaeger, 2013; Fine, Jaeger, Farmer, & Qian, 2013; Jaeger & Snider, 2013). Note that the 'speaker' to whom participants are adapting in our experiments is just the experiment itself. To more directly establish the presence of adaptation effects in the resolution of quantifier scope ambiguities, future studies can manipulate and compare the quantifier biases of different speakers, both on a separate per-quantifier bias and across many quantifiers. Another direction for future work is to compare the effects of primed and unprimed trials within-subject, giving the same participants ambiguous targets preceded both by irrelevant fillers as well as by primes that force one reading. This would provide separate measures of a participant's unprimed bias and average responding on prime trials. If the downward shift from the baseline condition to the average response in a primed condition is due to an adaptation effect that accumulates over the course of the experiment, the effect should be suppressed in a within-subject comparison. If there is a difference between baseline and primed trials, it would suggest that context effects depend on the immediately preceding trial (though not the prime scope).

While information about a given quantifier's bias is aggregated across trials within a single quantifier, we also find that information about the possible readings for EACH may affect the scope bias of EVERY and vice versa, suggesting that updating of scope bias may involve some shared semantic properties between just these two quantifiers. In contrast, the online assignment of relative scope (which draws on quantifier biases as one input, and which is reflected in the trial-by-trial priming effects) is not sensitive to these same features.

What semantic properties might EACH and EVERY share with each other, but not with ALL or the numbers? The semantic feature of 'distributivity' is most closely associated with the quantifiers EACH and EVERY. EACH has been analyzed as a lexicalized distributivity operator (see Roberts, 1987 for discussion), while EVERY, while still associated with distributivity, has been argued to be either optionally distributive (Beghelli and Stowell, 1997), or distributive over at least some, but perhaps not all atomic events in the set (Tunstall, 1998). We discuss the issue of distributivity further in the General Discussion of the main text.

Steedman's (2012) book offers another possible explanation. He proposes that only the denotations of the quantifiers EVERY and EACH contain the universal logical operator, ∀. Other quantifiers, like A, THREE and ALL, use entirely different formal combinatorial machinery (Skolem functions and constants; see also Reinhart, 1997; Winter, 2001). Along the continuum from lumping to splitting, Steedman's proposal lies in the middle. He groups EVERY and EACH into a single category, but splits these off from all the other quantifiers. As we mentioned, the absence of between-quantifier priming effects on a trial-by-trial basis gives no reason for grouping EACH with EVERY, or any other quantifiers together. If we assume, however, that the representational similarities posited by Steedman guide the rapid learning observed in the present study rather than online LF construction on a trial-by-trial basis, then Steedman's theory would predict context effects between the numbers and ALL, and between EVERY and EACH. We find one of these effects (the context effects between sentences with EVERY and EACH) but not the other (no effect between numbers and ALL). At first blush this suggests that Steedman's theory is right only in part. However, this theory leaves room for possible differences between the functions employed by various non-universal quantifiers. If these differences are sufficient to prevent generalization across quantifiers, then the context effects may be fully consistent with Steedman's proposal.

References

Beghelli, F., & Stowell, T. (1997). Distributivity and negation: The syntax of each and every. In Szcabolcsi, A. (Ed.), *Ways of scope taking* (pp. 71-107). Netherlands: Springer.

Champollion, L. (2010). *Parts of a whole: Distributivity as a bridge between aspect and measurement.* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia, PA.

Fine, A. B., Jaeger, T. F., Farmer, T. F., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. PLoS ONE, 8(10). doi:10.1371/ journal.pone.0077661

Fine, A. B., & Jaeger, T. F. (2013). Evidence for Implicit Learning in Syntactic Comprehension. Cognitive Science, 1-14. doi:10.1111/cogs.12022

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. Cognition, 127(1), 57-83. doi:10.1016/j.cognition.2012.10.013

Kamide, Y. (2012). Learning individual talkers' structural preferences. Cognition, 124(1), 66-71. doi:10.1016/j.cognition.2012.03.001

Reinhart, T. (1997). Quantifier scope: how labor is divided between QR and choice functions. *Linguistics and Philosophy*, *20*, 335–397.

Roberts, C. (1987). *Modal Subordination, Anaphora, and Distributivity.* Doctoral dissertation. University of Massachusetts at Amherst, MA.

Steedman, M. (2012). *Taking scope: The natural semantics of quantifiers*. Cambridge, MA: MIT Press.

Tunstall, S. L. (1998). *The interpretation of quantifiers: Semantics and processing.* (Unpublished doctoral dissertation), University of Massachusetts, Amherst, MA.

Winter, Y. (2001). *Flexibility principles in Boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. Boston MA: MIT press.

Yildrim, I., Degen, J., Tanenhaus, M. K., & Jaeger, T. F. (in press). Talker-specificity and adaptation in quantifier interpretation. Journal of Memory and Language.