An Experimental Note on Distributivity and Scope*

Manuel Križ · Mora Maldonado

Laboratoire de Sciences Cognitives et Psycholinguistique (ENS, EHESS, CNRS)
Département d'Études Cognitives, École Normale Supérieure, PSL Research University

Institut Jean-Nicod (ENS, EHESS, CNRS)
Département d'Études Cognitives, École Normale Supérieure, PSL Research University

August 7, 2018

Abstract

Experimental data are frequently used in formal semantics when intuitions about the existence of readings are shaky, and they can even help us to discover readings that go unnoticed by simple introspection (Marty et al., 2014). We report an experiment that tested the existence of distributive readings with plural noun phrases in object position (e.g. 'A mouse is painting all the penguins'), which has been the subject of some disagreement in the literature (Steedman 2012). Our findings indicate that these readings do, indeed, exist, and we will suggest that their comparative marginality can be explained without recourse to a ban on inserting the distributivity operator in derived scope positions.

1 Introduction

Sentences with a definite plural noun phrase in subject position, such as (1), are truth-conditionally compatible with, roughly speaking, three types of situations:[1]

(1) The mice are painting a penguin.
  a. Every mouse is painting one and the same penguin.
  b. Every mouse is painting a penguin, but it's not the same one for all of them.
  c. The mice are collaboratively painting a penguin (jointly making the painting, etc.).

It is not so clear that the situation is analogous for sentences involving plural definites in object position. While sentence (2) is certainly compatible with scenario (2a), its status in the *co-variation* scenario in (2b) is dubious.[2]

(2) A mouse is painting the penguins.

We are indebted to Emmanuel Chemla for valuable comments and discussion. The research leading to this work was supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 313610, and by ANR-10-IDEX-0001-02 PSL and ANR-10-LABX-0087 IEC.

[1] We are simplifying here by ignoring situations that are, as it were, in between the latter two, such as several teams of mice each collaboratively painting one and the same penguin.

[2] There is not analogue of situation (1c) here because every collective being-painted is also an individual being-painted, so that there is no genuinely collective reading with respect to the object position of *paint*.

a. Every penguin is being painted by a single mouse.
b. ??Every penguin is being painted by a mouse, but it's not the same for all of them.

The general goal of this paper is to explore experimentally the availability of different understandings for sentences such as (2), which involve a plural noun phrase in object position. Specifically, our aim is to assess whether or not these sentences can give rise to readings compatible with co-variation scenarios. Intuitions about such interpretations seem to be shaky, but experimental data can help us to discover readings that go unnoticed by simple introspection (Marty et al., 2014). Before introducing more details about the experimental set-up, we need to provide some formal semantics background on the ambiguities that arise with plural expressions.

1.1 The distributive/non-distributive contrast.

Most semantic approaches to plurality assume that sentences such as (1) can have two alternative readings. A first *non-distributive* interpretation is obtained *by default*, as soon as the predicate is applied to the plurality (see (3)).

(3) $\exists y.\text{penguin}(y) \wedge \textit{is}\text{-painting}(\iota x.\text{mice}'(x), y)$
There is a penguin, call it y, and a plurality made up of mice, call it X, such that X are painting y.

This reading is unquestionably true in the collective situation (1c). Whether it is also true in (1a) depends on what exactly one takes the predicate meaning to be. It is commonly assumed (Scha, 1981, a.m.o.) that if a predicate or relation is true of individuals, it is also true of the plurality made up of those individuals. Under this view, the reading in (3) is true both in (1a) and (1c), since it does not matter whether the individuals that make up the plurality are collaborating on the painting or are doing it independently. Crucially, this reading is not true in (1b): since the existential quantifier over penguins is the only quantifier in this reading, it has to be one and the same penguin — it contains only one quantificational operator, the existential, hence the penguin cannot vary.

To account for the reading that is compatible with the co-variation scenario in (1b), a second, *distributive* reading is assumed to come about through the insertion of an unpronounced distributivity operator D at the level of the verb phrase. This operator, which can be thought of as a silent version of the English word *each*, takes as its argument a predicate P and returns a new predicate that is true of a plurality if and only if P is true of all the atomic individuals that make up that plurality:

(4) The mice D are painting a penguin.
$\forall x. x \preceq_{AT} \iota z.\text{mice}(z) \rightarrow \exists y.\text{penguin}(y) \wedge \textit{is}\text{-painting}(x, y)$
For each atomic member x of the maximal plurality of mice, there is a penguin y such that x is painting y.

Since the existential quantifier *a penguin* is in the scope of the distributivity operator, which implicitly universally quantifies over atomic individuals, the penguin is allowed to vary by mouse (although it does not have to).

1.2 Plural quantifiers.

The observation from (1) holds not only with definite plurals, but with all plural DPs, including those headed by numerals and *all*. The sentences in (5), for example, are compatible with the same kinds of situations as (1); namely, with both *covariation* (where penguins vary by mouse) and *non-covariation* (where there is only one penguin) scenarios.

(5) a. Two mice are painting a penguin.
b. All the mice are painting a penguin.

It is natural to assume that the sentences in (5) are ambiguous in the same way as (1): as long as both *all* and numerals quantify over pluralities, (5a) and (5b) can receive their non-distributive reading by default, and their distributive reading — compatible with covariation scenarios — through insertion of the D operator. For instance, in (5b), the indefinite *two mice* asserts the existence of a plurality of two mice, of which the predicate *are paiting a penguin* can be true either non-distributively (6a) or distributively (6b), depending on whether or not the D operator has been inserted:[3]

(6) a. Two mice are painting a penguin.
$\exists x.\text{mice}(x) \wedge |x| = 2 \wedge \exists y.\text{penguin}(y) \wedge \text{painting'}(x, y)$
There is a plurality made of two mice X and a penguin y such that X are painting y.
b. Two mice D are painting a penguin.
$\exists x.\text{mice}(x) \wedge |x| = 2 \wedge D(\lambda z.\exists y.\text{penguin}(y) \wedge \text{painting'}(z, y))(x)$
There is a plurality made of two mice X such that, for each atomic part x of X, there is a penguin y such that x is painting y.

This distributive/non-distributive distinction is reminiscent of, but not identical to, the *scope ambiguities* that arise with double-quantified sentences such as (7).

(7) Every mouse is painting a penguin.
a. $\forall x.\text{mice}(x) \rightarrow \exists y.\text{penguin}(y) \wedge \text{painting}(x, y)$
For every mouse x, there is a (potentially different) penguin x such that y is painting x.
b. $\exists y.\text{penguin'}(y) \wedge \forall x.\text{mice'}(x) \rightarrow \text{painting}(x, y)$
There is a single penguin x such that for every mouse x, y is painting x.

Sentence (7) has two different readings, (7a) and (7b), depending on the relative scope of the quantifiers *every* and *a* (May, 1985; Fox, 2000, a.m.o). In (7a), the universal quantifier *every* takes wide scope above the existential, matching the linear order of the sentence (*surface-scope reading*). Instead, in (7b), the scope relation is not isomorphic to linear order: the existential takes wide-scope above the universal (*inverse-scope reading*). Surface-scope interpretations are considered to be basic, whereas inverse-scope readings are often assumed to be derived by applying a *scope-inverting* operation, which reverses the relative scope of the quantifiers at the interpretation stage.[4]

Like the quantified-sentences in (5), scope-ambiguous sentences like (7) are also compatible with *covariation* and *non-covariation* scenarios. One might thus come to conclude that this is really the same phenomenon, but there are good reasons to believe that this is not the case. One of the main arguments comes from differences in the semantics of plural and singular quantifiers (Champollion, 2010; Winter, 2001; Dowty et al., 1987). Unlike morphologically singular quantifiers (e.g. *every*), plural quantifiers are compatible with necessarily collective verb phrases[5].

[3] In the case of the universal *all*, we should note that, as far as truth conditions are concerned, *all the mice* has actually the same meaning as *the mice*. The semantic effect of *all* is of a different sort, such as adding some additional constraints on the predicates that noun phrase may combine with (Champollion, 2010), eliminating the possibility of non-maximal interpretations of the definite plural (Brisson, 1998; Križ, 2015), or removing logical trivalence effects (Križ, 2015).

[4] The need for a scope-inverting operation is uncontroversial, but there are different possible implementations: this operation has been viewed as syntactic movement (May, 1985; Fox, 2000; Montague, 1972), semantic type-shifiting, or a more complex mapping at the syntax-semantics interface (Steedman, 2012; Reinhart, 1997; Beghelli and Stowell, 1997).

[5] Indeed, the sentences in (5) can also have *collective* understandings, which do not exist for scope-ambiguous sentences with singular quantifiers like (7).

(8) a. All the mice are painting a penguin together.
b. #Every mouse is painting a penguin together.

(9) a. All the mice are gathering in the hallway.
b. #Every mouse is gathering in the hallway.

This suggests that singular quantifiers are *inherently* distributive, quantifying directly over atomic individuals, and therefore incompatible with collective predicates. Plural quantifiers, on the other hand, quantify over pluralities. As a result, the D-operator must be inserted to derive distributive interpretations compatible with *covariation* scenarios. For the sake of brevity, we will not further explore the details of this argument. The reader is referred to Winter (2001) and Champollion (2010) to a deeper discussion of these differences.

'Object distributivity': inverse-scope distributive readings. The existence of a mechanism such as the distributivity operator directly accounts for 'subject-distributivity': the predicate is distributed over each member of the plural subject, allowing objects to covary (e.g. (1) and (5)). It has been a matter of some debate whether the same mechanism can derive what we could call 'object-distributivity', namely, a reading that allows the *subject* to covary with each member of the plural *object* (for an overview, see Steedman, 2012: §3.1–3.3). Such reading would make the sentence (2) (repeated in (10)) compatible with a *covariation* scenario, where there is a different mouse painting each penguin (see scenario in (2b)).

The derivation of 'object-distributivity' would require applying the scope-inverting operation before inserting the D operator: the plural noun phrase should raise and the D operator should be inserted below its derived position, as it is schematically shown in (10a). The resulting truth-conditions are given in (10b).

(10) A mouse is painting the/all the/two penguins.
a. ??the/all the/two penguins D λx a mouse are painting x.
b. ??'Every penguin is being painted by (potentially varying) mice.'

Such *inverse-scope distributive readings* are difficult to access introspectively. This is quite different from the scope ambiguity of the sentence that is obtained by replacing the definite plural with a genuine universal quantifier: in (10), it is quite apparent that the sentence has a reading where *every penguin* takes scope of *a mouse* and therefore allows covariation of mice with penguins.

(11) A mouse is painting every penguin.
For every penguin y, there is a (potentially different) mouse x such that x is painting y.

The apparent introspective unavailability of *inverse-distributive readings* could have two different sources. It is possible that these readings are simply not generated by the grammar, for example, because the D-operator cannot be inserted in derived scope positions (Steedman, 2012). Alternatively, the apparent unavailability might be the reflex of processing cost or dispreference, resulting from the additive effect of applying two different interpretative operations during parsing (i.e. scope-inversion and insertion of D).

1.3 Previous psycholinguistic work.

The question of whether alternative readings of ambiguous sentences differ in terms of preference or cognitive cost has been a focus of attention in psycholinguistic research for some time. For distributive/non-distributive ambiguities, several studies have shown that distributive readings of sentences such as (1) tend to be both dispreferred and costly in comparison to non-distributive ones (Dotlacil, 2010; Kaup et al., 2002; Ussery, 1998; Syrett and Musolino, 2013; Musolino, 2009; Brooks and Braine, 1996; Brasoveanu and Dotlačil, 2015). Recent priming studies have also revealed an asymmetry between distributive and non-distributive readings (Maldonado et al., 2017), suggesting that these differences in parsing have a semantic source. These findings are well accounted for by the picture we described above, in which non-distributive readings are primitives, while distributive readings are derived through an additional operation (in this case, insertion of *D*).

The general preference for non-distributive readings was found for sentences involving a range of different plural expressions. However, it has been suggested that the degree to which distributive interpretations are dispreferred or costly might also be partially determined by lexical properties of the quantifier, among other factors (Dotlacil, 2010).This suggestion, however, mostly comes from a *between-studies* comparison, and it has not been directly tested.

Experimental literature on scope ambiguities has also suggested that scope preferences have a lexical component, such that quantifiers differ on the scope they are likely to take (wide vs. narrow; Ioup, 1995; Turnstall, 1998; Feiman and Snedeker, 2016). The universal quantifier *each*, for instance, seems to have a stronger preference for taking wide scope than the quantifier *every*[6]. These quantifier-specific scope preferences seem to be partially independent of whether the relative scope assignment is isomorphic to linear order. For example, many studies have found that *every* displays a general preference to take wide scope over *a* for both possible linear orders (Bott and Radó, 2009; Raffray and Pickering, 2010; Chemla and Bott, 2015; and see results and discussion in Chapter 6. The scope-inverting operation, however, might still have an effect on parsing, modulating quantifier-specific preferences (Anderson, 2004; Lidz and Musolino, 2002; Viau et al., 2010; see Pylkkänen and McElree, 2006 and Feiman and Snedeker, 2016 for discussion).

The status of *inverse-distributive readings* remains mostly unexplored in the literature, with the exception of Gil (1982). In an early study, Gil tested the availability of different readings of sentences involving numeric expressions (e.g. 'Two boys saw three girls'), and suggested that inverse-distributive readings are mostly unavailable for these sentences.

This paper aims at testing whether *inverse distributive readings* (i.e. 'object-distributivity') for sentences involving different plural expressions can be accessed by speakers. If these readings are marginally available, it would suggest a specific interaction between two distinct phenomena, distributivity and scope-inversion, which remains mostly unexplored experimentally. As a secondary goal, we aim to assess to which extent the availability of these readings might depend on the specific quantifier. The degree to which these readings are accepted might vary across plural quantifiers, depending on quantifier-specific scope and distributivity preferences. The outcome for different plural expressions might therefore shed light on their lexical properties.

2 Methods

2.1 Participants

We recruited 163 subjects via Amazon's Mechanical Turk, three of whom were excluded from the analysis for not being a native speaker of English. IPs were restricted to the United States and

[6]Some studies that compare quantifier-specific scope preferences have also included sentences such as (5a) and (5b) (Feiman and Snedeker, 2016), treating them as scope ambiguous sentences. As expected, the preferred reading for these cases is the non-distributive one, which in these studies was taken to be a preference for inverse-scope interpretations.

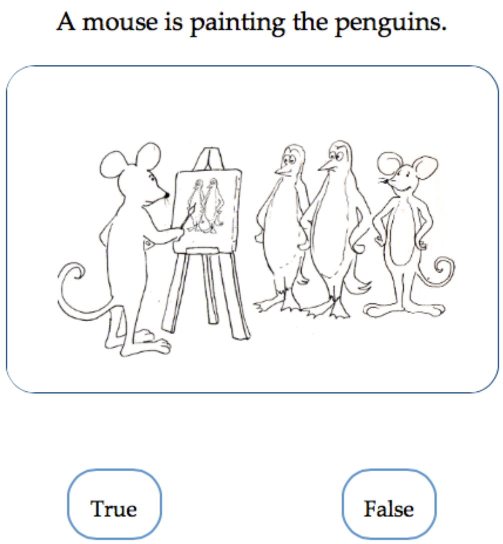

Figure 1: Example item from the INVERSE/THE/BOTH condition.

participants were paid 1.5 USD for their participation. The experiment itself made use of the Ibex platform (http://spellout.net/ibexfarm/).

2.2 Task and Design

Participants were instructed to perform a Truth-Value Judgment Task (TVJT). Each item consisted of a picture and a sentence (an example is given in Figure 1), and subjects had to judge the sentence as either *true* or *false* with respect to the picture.

The experimental design involved three fully crossed factors, FRAME, DETERMINER, and TRUTH, for a total of $2 \times 3 \times 3$ conditions. The sentences could follow one of two FRAMEs depending on whether the DP of interest was in subject (SURFACE frame) or object (INVERSE frame) position. The other argument position was always filled by a singular indefinite.

(12) a. DET <animals> ... an <animal>. (SURFACE)
b. An <animal> ... DET <animals> (INVERSE)

The DETERMINER could be either *all the*, plain *the*, or *two*. All verbs were transitive eventive predicates, and each predicate came associated with a subject and an object animal species (for example, when the predicate was *paint*, it was always mice doing the painting and penguins being painted).

All of these sentences have two potential readings, a distributive reading and a non-distributive one. In the SURFACE frame, the distributive reading is a regular instance of distributivity. The distributive reading in the INVERSE frame is the one whose existence is in question (e.g. *inverse distributive readings*). (13) shows the example of an INVERSE/ALL sentence with its two potential readings.

(13) A mouse is painting all the penguins.
a. Every penguin is being painted by a (possible different) mouse. (DIST)
b. The same one mouse is painting all the penguins. (BOTH)

The factor TRUTH determined which of these readings was true given the picture. In the

NEITHER conditions, neither reading was true; in the DIST conditions, only the distributive reading was true; and in the BOTH conditions, both readings were true. Note that the pictures in the latter conditions were non-distributive, but also non-collective (no collaborative painting, see (1a)). The distributive reading was therefore true because "one and the same mouse" is just a special case of "a (possibly different) mouse".

There were three types of pictures, and the factors FRAME and TRUTH together determined which one was instantiated: A_1VB_n, A_nVB_1, and A_nVB_n (Table 1). Here, A is to be understood as the restrictor noun of the subject noun phrase, B as the restrictor noun of the object noun phrase, and V as the transitive predicate. A picture of type A_1VB_n had one A standing in relation V to n Bs, and an additional A that did not stand in the relation to anyone. Pictures of type A_nVB_n had n As and n Bs in a one-to-one V relation. Pictures of type A_nVB_1 had n As all standing in relation V to the same one B, and an additional B that did not stand in the relation to anyone.[7] When the determiner was ALL, then n was set to 3, otherwise it was 2. The reason for this was to fulfil the *more than two* antipresupposition for all' (Heim, 1991; Percus, 2006). Examples of the pictures for a $n = 2$ are shown in Figure 2. Table 1 shows how FRAME and TRUTH determined a picture type.

img/

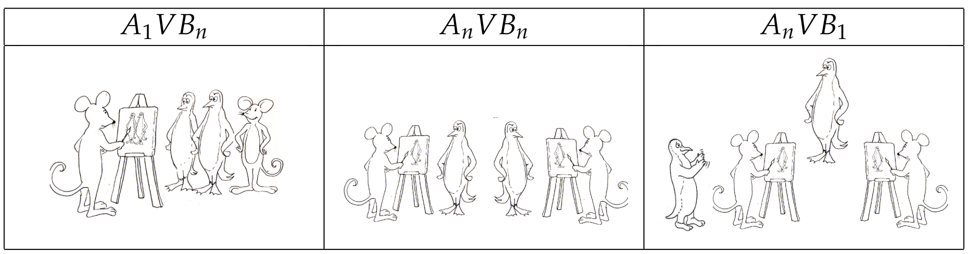

Figure 2: Picture types with example for $n = 2$.

FRAME	SENTENCE	TRUTH		
		NEITHER	DIST	BOTH
SURFACE	All the *A*s *V* a *B*. The *A*s *V* a *B*. Two *A*s *V* a *B*.	A_1VB_n	A_nVB_n	A_nVB_1
INVERSE	An *A* *V* all the *B*s. An *A* *V* the *B*s. An *A* *V* two *B*s.	A_nVB_1	A_nVB_n	A_1VB_n

Table 1: Determination of picture type by FRAME and TRUTH.

Each subject saw five instances of each condition (with different verbs and animal species) for a total of 60 items per participant. Materials, data and analyses are available online in this link.

3 Results

Percentage of 'true' responses for all values of DETERMINER, FRAME, and TRUTH are shown in Figure 3. Our analyses were performed under two basic considerations. First, to the extent a

[7]The additional bystander A in the A_1VB_n pictures served the purpose of satisfying the plurality presupposition of *(all) the As* in the SURFACE conditions, and analogously the additional B in the A_nVB_1 pictures.

distributive reading does not exist, *dist* conditions (in blue) should behave like *neither* conditions (in orange) within a frame. Second, if inverse-distributive readings are less available than basic distributive readings, the *dist* condition should be lower in the inverse than in the surface frame. These comparisons form the basis of our analyses.

Now, given that we are comparing three different noun phrases, we wish to ask the question of "Which, if any, of *all*, *the* and *two* behave the same?". A null-hypothesis testing framework does not allow one to directly address a question of this form. For this reason, we decided to perform Bayesian analyses of the data. We used STAN/`rstanarm` to fit a variety of mixed-effects logit models and compared these models with or without quantifier-specific factors via leave-one-out cross-validation (Vehtari et al., 2016). For the sake of brevity, the details of our analyses are provided in the Supplementary Results section (Section 6). The general results of our analyses are the following:

1. Distributive readings are clearly more available with *all* than with the other quantifiers ($\Delta_{elpd} = -83.6$ in *surface* and $\Delta_{elpd} = -11$ in *inverse* frame).

2. There is negligible evidence for a difference between *the* and *two* in the availability of distributive readings ($\Delta_{elpd} = -0.4$ in *surface* and $\Delta_{elpd} = -3.0$ in *inverse* frame, with the effects going in opposite directions).

3. Evidence for inverse scope distributive readings with all quantifiers ($\Delta_{elpd} = -165.7$ for *all*, $-37.$ for *the*, -8.1 for *two*).

4. Assuming that the availability of inverse scope distributive readings is modulated by (i) the quantifier's inherent preference for distributive readings (as in 1. above) and (ii) the quantifier's scope preferences, we find models ranked as follows with respect to quantifiers' propensity for inverse scope: *the* > *two* > *all* is better than *the* > *all* = *two* ($\Delta_{\text{elpd}} = -4.9$, $se(\Delta_{\text{elpd}}) = 3.5$), which is better than *the* = *two* > *all* ($\Delta_{\text{elpd}} = -3.0$, $se(\Delta_{\text{elpd}}) = 6.9$). The strength of the evidence here is not overwhelming, and so the picture is overall somewhat muddy, but it is worth pointing out the following: among these three determiners, *all* actually has the lowest propensity for inverse scope, despite the fact that inverse-scope distributive readings are most easily available with *all*.

   Note also that it is clear that idiosyncratic scope preferences do exist. A model with only an across-the-board factor for inverse scope that was not quantifier-specific was markedly worse than the next best model ($\Delta_{\text{elpd}} = 22.2$, $se(\Delta_{\text{elpd}}) = 6.9$).

4 Discussion

We have provided counter-evidence to the common assumption that inverse scope distributive readings with plural quantifiers do not exist. These readings are shown to be marginally available for *all*, definite descriptions and numerals, suggesting that it is possible to insert the distributivity operator in derived-scope positions.

The marginal status of these readings can be understood as resulting from the interaction of distributivity and scope-inversion. That is, low acceptability rates for inverse distributive readings could be indicative of dispreference or high processing cost (for this particular reading), and these would in turn be caused by the combination of two different interpretative mechanisms. The degree to which inverse distributive readings are available, however, also seems to be modulated by quantifier-specific preferences for both distributivity and scope.

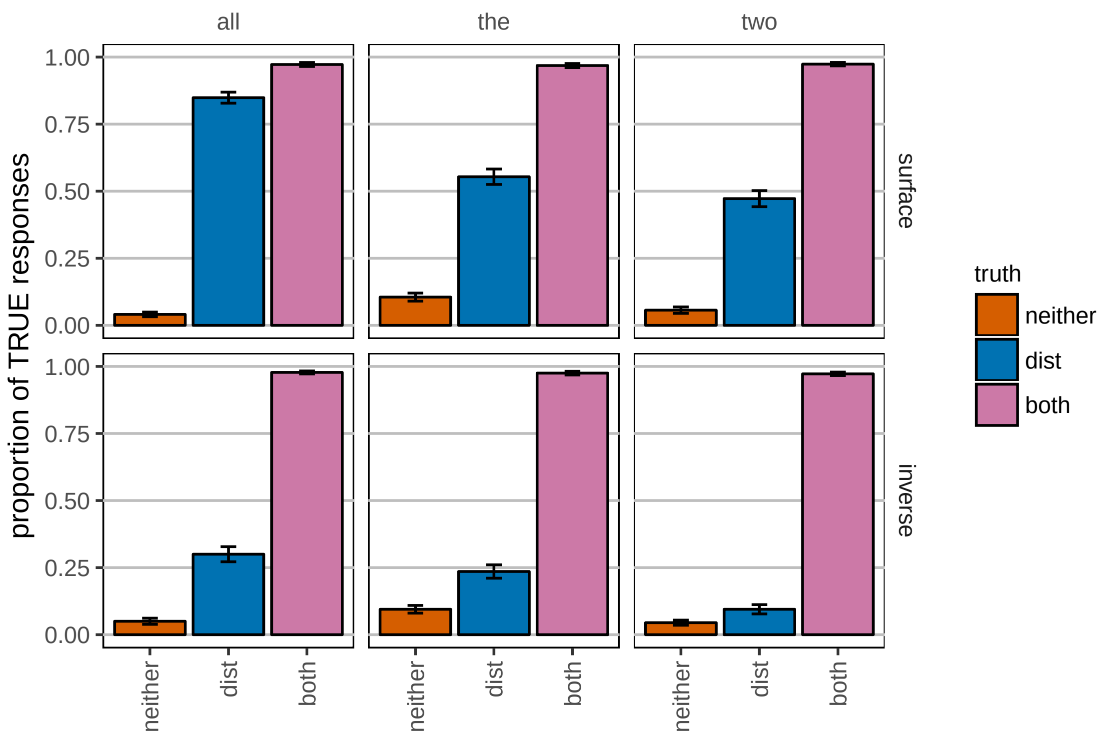

Figure 3: Mean proportions of *true* responses in all conditions. Error bars correspond to standard error of the mean per participant.

First, we found evidence for lexically conditioned differences in the availability of distributive readings. While non-distributive readings are universally accepted[8], distributive interpretations are in general more readily available with *all* than with definites and numerals. These findings pattern with previous results in the literature, where the relative preference for distributive readings was found to be higher for *all* than for numerals (Feiman and Snedeker, 2016). On the assumption that distributive readings are always derived by applying the *D* operator, the difference between determiners is presumably grounded in a frequency effect: the probability that the silent *D* operator is there might be higher after *all* than after *the* and *two*, making the parsing with *D* more likely in the former case.

Second, we also found some evidence for a difference in the propensity of plural noun phrases to take inverse scope. Noun phrases headed by *the* are the most likely ones to take inverse-scope. On the other end of the hierarchy, noun phrases with *all* have the lowest propensity for inverse scope. Interestingly, quantifier-specific scope preferences of this sort have been argued to exist for other quantifiers (Ioup, 1995; Turnstall, 1998; Feiman and Snedeker, 2016).

Overall, *all* is still the determiner with which inverse-scope distributive readings are most available, but that is purely because its preference for distributive readings is much stronger than with other determiners, and it is just attenuated, but not nullified, by the greater dispreference for inverse scope. In other words, *all* has a higher preference for distributivity than a dispreference for inverse scope.

Taken together, our findings suggest not only that scope and distributivity do interact — giving

[8] Let us note as a secondary point that the fact that non-distributive readings are fully available supports the assumption of closure under fusion for lexical predicates, which predicts that the non-distributive reading is made true equally by collective and separate individual actions ((1a) and (1c) from the beginning of this article).

rise to inverse-distributive readings — , but also that the rate at which these phenomena occur is at least partly dependent on specific biases of individual quantifiers.[9] The fact that quantifiers' preferences play a role in the availability of distributive and inverse-scope interpretations has practical implications for both experimental and intuition-based approaches to these phenomena, as general claims about distributivity and scope should be tested for a range of different determiners.

Before concluding, let us note an alternative explanation of our findings, whereby the asymmetry in the availability of inverse distributive readings between *all*, on the one hand, and *the* and *two*, on the other, has a different source than a lexical preference. Naturally, this would independent of the fact that inverse distributive readings do exist for the three plural noun phrases —i.e. they are also marginally available for definites and numerals.

One could imagine, for instance, that *all* obtains its distributive readings in a different way from definite plurals and numerals. The most natural form this view could take is to assign to *all* its regular universal quantifier meaning and assume that there is a collectivising operator that can be applied to it to obtain collective readings (see (8) and (9)).[10] *All* is then, in a sense, the mirror image of definite plurals: There is no operator in distributive readings, but there is one in non-distributive readings — exactly the opposite of what is assumed for definite plurals. Distributive interpretations would then be obtained by default for *all*, explaining why these readings are more available than for the other determiners (in both surface and inverse frames).[11] Potential counterarguments to this view come from the fact that *all* does not consistently contrast with *two* and *the*, as observed for, for instance, scope preferences.

5 Conclusions

This paper tested the existence of *inverse distributive readings* for sentences involving plural noun phrases in object position. By using a range of plural quantifiers, we additionally explored how lexical preferences might modulate the propensity to take distributive and inverse-scope interpretations. Our findings make two main contributions. First, we found that, despite being quite marginal, inverse-distributive readings exist for *all*, definite descriptions and numerals. Moreover, the availability of these readings is shown to be partly conditioned by quantifier' preferences regarding both distributivity and scope.

6 Supplement on statistical analysis

Since we are comparing three types of noun phrases, the questions we wish to ask all have the fundamental form "Which, if any, of these three are the same?". A null-hypothesis testing framework does not allow one to straightforwardly address a question of this form. For this reason,

[9] Some authors have suggested that differences in quantifiers' scope-taking behaviour might not reflect a simple frequency effect, but a deeper difference in the nature of the scope-inversion operation (Feiman and Snedeker, 2016). Instead of positing a unique scope-inversion operation for all quantifiers, these approaches have suggested that quantifiers might differ from each other in their scopal mechanism (in the lines of Beghelli and Stowell (1997) and Steedman (2012)). At the moment, however, there is no enough evidence to tease these options apart. Given that all the noun phrases tested in this paper can take inverse scope, we maintain the general view of a single mechanism for all quantifiers.

[10] Such a picture is argued for by Winter (2001) for independent reasons — his system also predicts inverse scope distributive readings for definite plurals and numerals, but does so via a different mechanism that is independent of the meaning of *all*.

[11] The still high rate of non-distributive readings for *all* in the surface scope position could be understood by positing that there are multiple ways to arrive at such non-distributive interpretation.

among others, we opted to conduct a Bayesian analysis. Mixed-effects logit models were fitted with STAN using the `rstanarm` library. The prior distribution for all parameters was Student's t distribution with mean 0 and 5 degrees of freedom.[12] 10,000 samples of the likelihood of each data point were drawn after 10,000 burn-in iterations from four chains, for a total of 40,000 samples. Models were evaluated by leave-one-out cross-validation, approximated by Pareto-smoothed importance sampling with the `loo` package (Vehtari et al. 2016) based on the resulting 40,000 posterior samples.

Central to our analysis are the following two considerations.

- To the extent that distributive readings do not exist, the DIST conditions (in a given frame) should behave like the NEITHER conditions.
- To the extent that acceptance in the DIST/INV condition is lower than in the DIST

Our analysis was built on the following two considerations. First, to the extent that inverse-scope distributive readings do not exist, the DIST/INV conditions should behave like the NEITHER conditions for the same determiner. Second,

In order to avoid having to fit an excessive number of models, our analysis proceeded in steps corresponding to a number of questions that are interesting to ask about our data.

6.1 Analysis 1: Baseline

For each determiner, the *neither* conditions serve as a baseline, in that to the extent that distributive readings don't exist with that determiner, the DIST conditions should look just like the NEITHER conditions.

To reduce the number of models to be computed and compared later on, we first restricted our data to the NEITHER conditions (from both frames, since there is no reason to expect a difference between the frames here) and performed inference on whether the baseline is the same for all three determiners.

In particular, we fitted mixed effects logit models of the following general form:

$$Y_{isd} \sim \mathsf{bernoulli}(\mathsf{logit}^{-1}(\pi_{isd}))$$
$$\text{with: } \pi_{isd} = \alpha_d + u_{1s}$$

where s was the subject and d the determiner. Five models were fitted which differed in which, if any, of the α-parameters were the same. Results of the leave-one-out cross-validation are shown in Table 2. The model where *all* and *two* are the same, but *the* is different, is roughly equivalent to the model where all three are different ($\Delta_{\text{elpd}} = -0.6$, $se(\Delta_{\text{elpd}}) = 0.7$), and there is substantial evidence in its favour compared to the next best model, where *the* and *two* are equal and *all* is different ($\Delta_{\text{elpd}} = -11.7$, $se(\Delta_{\text{elpd}}) = 5.3$).

This is in accordance with theoretical expectations: sentences with *all* and *two* are simply false in the NEITHER conditions, whereas those with *the* suffer a homogeneity violation rather than being simply false. The latter would plausibly translate to a slightly higher rate of acceptance (Schwarz 2013), but there is little reason a priori to expect a difference between *all* and *two*.

To simplify matters, we therefore initially adopted the assumption that $\alpha_{\text{all}} = \alpha_{\text{two}}$ as the basis of our further analyses. A later batch of models where this assumption was dropped confirmed the results described here.

[12] We spot-checked by fitting a few models with slightly different prior distributions and found our results did not change.

Parameters	elpd	*se*(elpd)
$\alpha_{\text{the}}, \alpha_{\text{all}} = \alpha_{\text{two}}$	−1072.9	41.2
$\alpha_{\text{the}}, \alpha_{\text{all}}, \alpha_{\text{two}}$	−1073.5	41.2
$\alpha_{\text{the}} = \alpha_{\text{two}}, \alpha_{\text{all}}$	−1084.6	41.5
$\alpha_{\text{the}} = \alpha_{\text{all}}, \alpha_{\text{two}}$	−1088.6	41.7
$\alpha_{\text{the}} = \alpha_{\text{all}} = \alpha_{\text{two}}$	−1090.5	41.9

Table 2: Estimated log predictive likelihood for models in Analysis 1.

6.2 Analysis 2: Distributivity preferences

This analysis is concerned with the availability of distributive readings, and the extent to which this differs between determiners.

6.2.1 Analysis 2a: Surface frame

Parameters	elpd	*se*(elpd)
$\beta_{\text{all}}, \beta_{\text{the}} = \beta_{\text{two}}$	−1698.2	41.4
$\beta_{\text{all}}, \beta_{\text{the}}, \beta_{\text{two}}$	−1698.6	41.3
$\beta_{\text{all}} = \beta_{\text{the}}, \beta_{\text{two}}$	−1781.8	42.3
$\beta_{\text{all}} = \beta_{\text{two}}, \beta_{\text{the}}$	−1884.5	39.0
$\beta_{\text{all}} = \beta_{\text{the}} = \beta_{\text{two}}$	−1907.9	39.5

Table 3: Estimated log predictive likelihood for models in Analysis 2a.

We first we restricted the data set to items from the SURFACE frame in the NEITHER or DIST conditions (thus excluding BOTH conditions, which were supposed to be, and were indeed, at ceiling) and fitted mixed effects logit models of the following general form:

$$Y_{isd} \sim \mathsf{bernoulli}(\mathsf{logit}^{-1}(\pi_{isd}))$$
$$\text{with: } \pi_{isd} = \alpha_d + u_{1s} + (\beta_d + u_{2s}) \cdot \text{DIST}_i$$

where s was the subject and d the determiner. DIST was 0 or 1 depending on whether the item was from the NEITHER or the DIST condition. As per Analysis 1, α_{all} and α_{two} were constrained to be identical.

The β-parameters are a measure of the preference of the distributive reading relative to the non-distributive reading. We fitted five models which varied in which, if any, of the three β-parameters were set to be identical. The results we obtained lead to two conclusions:

- Distributive readings are beyond doubt more preferred with *all* than with the other two types of DPs ($\Delta_{\text{elpd}} = -83.6$, $se(\Delta_{\text{elpd}}) = 12.8$).
- There is no evidence for a difference in the availability of distributive readings between the definite plural and the numeral ($\Delta_{\text{elpd}} = -0.4$, $se(\Delta_{\text{elpd}}) = 1.4$).

6.2.2 Analysis 2b: Inverse frame

This analysis mirrors that of the previous section, but applied to the INVERSE frame. The results are shown in Table 4.

Parameters	elpd	se(elpd)
$\beta_{\text{all}}, \beta_{\text{the}}, \beta_{\text{two}}$	−1421.2	41.2
$\beta_{\text{all}}, \beta_{\text{the}} = \beta_{\text{two}}$	−1424.3	41.2
$\beta_{\text{all}} = \beta_{\text{the}}, \beta_{\text{two}}$	−1435.3	41.4
$\beta_{\text{all}} = \beta_{\text{two}}, \beta_{\text{the}}$	−1494.9	41.2
$\beta_{\text{all}} = \beta_{\text{the}} = \beta_{\text{two}}$	−1496.3	41.2

Table 4: Estimated log predictive likelihood for models in Analysis 2b.

- There is dubious evidence for a difference in the availability of inverse-scope distributive readings for all three quantifiers ($\Delta_{\text{elpd}} = -3.0$, $se(\Delta_{\text{elpd}}) = 3.1$)
- If only one quantifier is different from the other two, then it is probably *all*, as in the surface conditions ($\Delta_{\text{elpd}} = -11.0$, $se(\Delta_{\text{elpd}}) = 8.1$).

6.2.3 Analysis 2c: Inverse-distributive readings

Parameters	elpd	se(elpd)
$\beta_{\text{all}}, \beta_{\text{the}}, \beta_{\text{two}}$	−1476.3	41.5
$\beta_{\text{all}}, \beta_{\text{the}} = \beta_{\text{two}}$	−1479.0	41.5
$\beta_{\text{all}}, \beta_{\text{the}}$	−1487.1	41.3
$\beta_{\text{all}}, \beta_{\text{two}}$	−1516.9	41.1
β_{all}	−1527.9	40.9
$\beta_{\text{the}}, \beta_{\text{two}}$	−1644.7	41.2

Table 5: Estimated log predictive likelihood for models in Analysis 2c. β-parameters set to 0 are omitted.

The analysis in the previous section targets differences in the availability of inverse-scope distributive readings between determiners, assuming the possibility of such readings. We therefore still need to establish whether, and for which determiners, such readings do, in fact, exist.

We can take the best model or models from the previous section and selectively set the β-parameters to 0 for certain determiners. To the extent that those models do worse than the original ones, we have evidence for the genuine existence of inverse-scope distributive readings.[13]

Table 5 shows the estimated log predictive likelihoods for various models. We find strong evidence that inverse-scope distributive readings do, indeed, exist for all three determiners, with the evidence being weakest, but still not negligible, for the case of *two* ($\Delta_{\text{elpd}} = -8.1$, $se(\Delta_{\text{elpd}}) = 6.9$).

6.3 Analysis 3: Scope-taking preferences

The availability of inverse-scope distributive readings for a determiner can be expected to be modulated by two factors: the availability of distributive readings with that determiner in general, and the propensity of the DP type to take inverse scope. Having established that the preference for distributive readings is to some extent determiner-specific even among plural DPs, we may now ask the same question about scope-taking behaviour: are any of *all*, *the*, and *two* more ready to take inverse scope than others?

[13]Note that we cannot include a subject intercept in these models as it would absorb the effect in a model with a β-parameter set to 0 and thereby make it appear as good as one where the β-parameter is not set to 0.

Parameters	elpd	*se*(elpd)
$\gamma_{\text{the}}, \gamma_{\text{all}}, \gamma_{\text{two}}$	−3072.2	57.6
$\gamma_{\text{the}}, \gamma_{\text{all}} = \gamma_{\text{two}}$	−3076.8	57.6
$\gamma_{\text{all}}, \gamma_{\text{the}} = \gamma_{\text{two}}$	−3081.4	57.6
$\gamma_{\text{all}} = \gamma_{\text{the}} = \gamma_{\text{two}}$	−3103.6	57.3
$\gamma_{\text{all}} = \gamma_{\text{the}}, \gamma_{\text{two}}$	−3104.3	57.4

Table 6: Estimated log predictive likelihood for models in Analysis 3 where $\beta_{\text{the}} = \beta_{\text{two}}$.

Parameters	elpd	*se*(elpd)
$\gamma_{\text{the}}, \gamma_{\text{all}}, \gamma_{\text{two}}$	−3071.8	57.5
$\gamma_{\text{the}}, \gamma_{\text{all}} = \gamma_{\text{two}}$	−3075.2	57.5
$\gamma_{\text{all}}, \gamma_{\text{the}} = \gamma_{\text{two}}$	−3082.4	57.6
$\gamma_{\text{all}} = \gamma_{\text{the}} = \gamma_{\text{two}}$	−3104.9	57.3
$\gamma_{\text{all}} = \gamma_{\text{the}}, \gamma_{\text{two}}$	−3105.2	57.4

Table 7: Estimated log predictive likelihood for models in Analysis 3 without constraints on β_{the} and β_{two}.

We thus fitted models of the following form on the full data set minus the BOTH conditions. Again there were five models, varying by which, if any, of the three γ-parameters were constrained to be identical.

$$Y_{isd} \sim \mathsf{bernoulli}(\mathsf{logit}^{-1}(\pi_{isd}))$$
$$\text{with: } \pi_{isd} = \alpha_d + u_{1s} + (\beta_d + u_{2s}) \cdot \text{DIST}_i + (\gamma_d + u_{3s}) \cdot \text{DIST}_i \cdot \text{INV}_i$$

where $\alpha_{\text{all}} = \alpha_{\text{two}}$ (conclusion from Analysis 1). We fitted both models where $\beta_{\text{the}} = \beta_{\text{two}}$, as concluded from Analysis 2a, and models where this constraint was lifted. As can be seen from Tables 6 and 7, there was no appreciable difference between these two sets of models.

It is clear that idiosyncratic scope preferences do exist among the three determiners we investigated, as the model without idiosyncratic differences is notable worse than the next best model ($\Delta_{\text{elpd}} = 21.5$, $se(\Delta_{\text{elpd}}) = 6.9$). Beyond that, the picture that emerges is not entirely decisive, with possible pictures ranked as follows: *the* > *two* > *all* is somewhat better than *the* > *all* = *two* ($\Delta_{\text{elpd}} = -4.9$, $se(\Delta_{\text{elpd}}) = 3.5$), which in turn is somewhat better than *the* = *two* > *all* ($\Delta_{\text{elpd}} = -3.0$, $se(\Delta_{\text{elpd}}) = 6.9$).

It is worth noting that despite the fact that inverse-scope distributive readings are most available with *all*, it is actually the determiner with the greatest aversion to taking inverse scope: the availability of inverse-scope distributive readings is higher than with the other determiners purely because distributive readings in general are more readily available with *all*.

References

Catherine Anderson. *The structure and real-time comprehension of quantifier scope ambiguity*. PhD thesis, Northwestern University Evanston, IL, 2004.

Filippo Beghelli and Tim Stowell. Distributivity and negation: The syntax of each and every. In *Ways of scope taking*, pages 71–107. Springer, 1997.

Oliver Bott and Janina Radó. How to provide exactly one interpretation for every sentence, or what eye movements reveal about quantifier scope. In *The fruits of empirical linguistics*, volume 1, pages 1–24. 2009.

Adrian Brasoveanu and Jakub Dotlačil. Processing pluralities: Syntax and the lexicon. *CUNY 28*, 2015. URL https://people.ucsc.edu/~abrsvn/CUNY_poster_2015.pdf.

Christine M Brisson. *Distributivity, maximality, and floating quantifiers*. PhD thesis, Citeseer, 1998.

Patricia J Brooks and Martin DS Braine. What do children know about the universal quantifiers all and each? *Cognition*, 60(3):235–268, 1996.

Lucas Champollion. Parts of a Whole: Distributivity as Bridge between Aspect and Measurement. pages 1–10, 2010. ISSN 04194209.

Emmanuel Chemla and Lewis Bott. Using structural priming to study scopal representations and operations. *Linguistic Inquiry*, 2015.

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

David Dowty et al. Collective predicates, distributive predicates and all. In *Proceedings of the 3rd ESCOL*, pages 97–115. (Eastern States Conference on Linguistics), Ohio State University Ohio, 1987.

Roman Feiman and Jesse Snedeker. The logic in language: How all quantifiers are alike, but each quantifier is different. *Cognitive Psychology*, 87:29–52, 2016.

Danny Fox. *Economy and semantic interpretation*, volume 35. MIT press, 2000.

David Gil. Quantifier scope, linguistic variation, and natural language semantics. *Linguistics and Philosophy*, 5(4):421–472, 1982.

Irene Heim. Articles and definiteness. *Semantics. An international handbook of contemporary research. Berlin: De Gruyter*, 1991.

Georgette Ioup. Some universals for quantifier scope. In J. Kimball, editor, *Syntax and Semantics 4*, pages 37–58. Academic Press, New York, 1995.

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.

Manuel Križ. *Aspects of Homogeneity in the Semantics of Natural Language*. PhD thesis, University of Vienna, 2015.

Jeffrey Lidz and Julien Musolino. Children's command of quantification. *Cognition*, 84(2):113–154, 2002.

Mora Maldonado, Emmanuel Chemla, and Benjamin Spector. Priming plural ambiguities. *Journal of Memory and Language*, 95:89 – 101, 2017. ISSN 0749-596X. doi: http://dx.doi.org/10.1016/j.jml.2017.02.002. URL http://www.sciencedirect.com/science/article/pii/S0749596X17300098.

Paul Marty, Emmanuel Chemla, and Benjamin Spector. Phantom readings: the case of modified numerals. *Language, Cognition and Neuroscience*, 30(4):462–477, 2014.

Robert May. *Logical Form: Its structure and derivation*, volume 12. MIT press, 1985.

Richard Montague. Pragmatics and intensional logic. In *Semantics of natural language*, pages 142–168. Springer, 1972.

Julien Musolino. The logical syntax of number words: Theory, acquisition and processing. *Cognition*, 111(1):24–45, 2009.

Orin Percus. Antipresuppositions. *Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as an empirical science*, 52:73, 2006.

Liina Pylkkänen and Brian McElree. The syntax-semantics interface: On-line composition of sentence meaning. In *Handbook of Psycholinguistics (Second Edition)*, pages 539–579. Elsevier, 2006.

Claudine N Raffray and Martin J Pickering. How do people construct logical form during language comprehension? *Psychological Science*, 21:1090–1097, 2010. ISSN 0956-7976. doi: 10.1177/0956797610375446.

Tanya Reinhart. Quantifier scope: How labor is divided between qr and choice functions. *Linguistics and philosophy*, 20(4):335–397, 1997.

Remko Scha. Distributive, collective and cumulative quantification. In J. Groenendijk, T. Janssen, and M. Stokhof, editors, *Formal Methods in the Study of Language*, volume 2, pages 483–512. Amsterdam: Mathematisch Centrum, 1981.

Florian Schwarz. Maximality and definite plurals – experimental evidence. In Emmanuel Chemla, Vincent Homer, and Grégoire Winterstein, editors, *Proceedings of Sinn und Bedeutung 17*, pages 509–526, 2013.

Mark Steedman. *Taking Scope*. MIT Press, 2012.

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

Susanne Lynn Turnstall. *The interpretation of quantifiers: semantics and processing*. PhD thesis, University of Massachusetts, Amherst, 1998.

Cherlon Ussery. Processing plural dps: Collective, cumulative, and distributive interpretations. University of Massachusetts, Amherst., 1998.

Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Manuscript, HIIT and Columbia University, 2016.

Joshua Viau, Jeffrey Lidz, and Julien Musolino. Priming of abstract logical representations in 4-year-olds. *Language Acquisition*, 17(1-2):26–50, 2010.

Yoad Winter. *Flexibility Principles in Boolean Semantics*. MIT Press, Cambridge, MA, 2001.