Existential and universal readings of pronouns across binary connectives: an experimental investigation [*†]

Keny Chatain[1], Benjamin Spector[1], and Nina Gregorio[2]

[1]Institut Jean-Nicod (ENS - EHESS - CNRS), Département d'Etudes Cognitives, Ecole Normale Supérieure, Paris, France; PSL Research University.
[2]University of Edinburgh

Abstract

Indefinites may co-vary with pronouns they do not c-command. The donkey configuration is the best-studied of these configurations. It is well-known that these configurations give rise to both existential and universal truth-conditions and that they don't require uniqueness (Heim, 1982). While there is a rich literature on pronoun readings in connection to donkey sentences (Champollion et al., 2017; Chierchia, 1992; Kanazawa, 1994, a.o.), similar questions also arise in simpler configurations without quantifiers, like cross-conjunction anaphora, and cross-disjunction anaphora (i.e. bathroom sentences, Roberts (1987)). In the case of bathroom sentences, it is still an open empirical question whether these sentences receive an existential reading (Elliott, 2020), a universal reading (Krahmer and Muskens, 1995) or both, and whether they require uniqueness (Gotham, 2019). In this paper, we present a series of truth-value judgment tasks aimed at probing the truth-conditions accessed in cross-conjunction and cross-disjunction configurations. We conclude that there is evidence that cross-disjunction sentences are

[*]Acknowledgments to be added…

[†]KC and BS conceived of the initial experiment together. They supervised NG, who created the materials and code for running the experiment. Together, all three authors revised the initial designs. KC wrote the manuscript, receiving comments from BS and NG.

ambiguous between an existential reading and a universal reading and that there is only evidence for existential readings in cross-conjunction sentences. We argue that this pattern is not predicted by any existing approach and discuss possible amendments. While the amendments are successful in predicting the patterns we observe, they rely on arbitrary and unmotivated stipulations. An additional consequence of our results is to show parallels between the readings obtained in donkey (i.e. quantified) sentences and the readings obtained in non-quantified sentences, suggesting a unified approach for both is desirable.

Contents

1 Introduction

Pronouns may co-vary with indefinites that do not c-command them. The most well-studied case is the case of *donkey sentences* in (1).

(1) Every farmer who owns a donkey$_i$ cherishes it$_i$.

Deriving the possibility of co-varying readings in such non-c-commanding configurations is a hard problem that has often required sophisticated analyses (Cooper, 1979; Evans, 1977, 1980; Heim, 1982; Kamp et al., 2011; van der Does, 1993). Even when an analysis does predict the possibility of co-variation, there remains the question of which truth-conditions ought to be predicted. Two main types of readings have been considered: the existential reading, paraphrased in (2a), and the universal reading, paraphrased in (2b).

(2) **Paraphrase:** *every farmer who owns a donkey cherishes…*

a. … *at least one of the donkey they own* (∃ reading)

b. … *all of the donkeys they own* (∀ reading)

For a sentence like (1), speakers sometimes access an existential interpretation and sometimes a universal one. There is evidence for this in both the theoretical literature (Champollion et al., 2017; Chierchia, 2009; Cooper, 1979; Kanazawa, 1994; Schubert and Pelletier, 1989) and the experimental literature (Denić and Sudo, 2022; Foppolo, 2008; Geurts, 2002; Sun et al., 2020). It is controversial whether these two interpretations genuinely constitute readings, whether they correspond to two different resolutions of under-specified truth-conditions Champollion et al. (2019) or whether the universal interpretation is obtained from the existential interpretation through defeasible pragmatic strengthening r.cite Elliott. The rest of this paper is agnostic about the nature of these interpretations and we refer to them for simplicity as readings.

The literature has also investigated variants of the donkey sentence in (1) with other quantifiers like *no*, as in (3), or *some*, as in (4). There, the evidence for ambiguity is

much scarcer. Experimental work has only found evidence for an existential reading, not for a universal reading. In theoretical work (Chierchia (2009), where the example is attributed to Kanazawa (1994)), the evidence for such readings mainly comes from *umbrella* sentences, like (5).

(3) No farmer who has a donkey$_i$ cherishes it$_i$

a. … *cherishes any of the donkeys they own* (∃ reading)

b. … *cherishes all of the donkeys they own* (∀ reading)

(4) Some farmer who has a donkey$_i$ cherishes it$_i$

a. … *cherishes some of the donkeys they own* (∃ reading)

b. … *cherishes all of the donkeys they own* (∀ reading)

(5) No person who has an umbrella left it at home today.

a. # …*left some of their umbrellas at home* (∃ reading)

b. ✓ …*left all of their umbrellas at home* (∀ reading)

In summary, there has been a reasonable amount of work on readings of donkey sentences. By contrast, much less attention has been given to the readings of indefinite-anteceded pronouns outside of the donkey configuration. Such configurations include cross-conjunction anaphora, as in (6a), and bathroom sentences (i.e. cross-disjunction anaphora), as in (6b).

(6) a. Sam owns a donkey$_i$ and it$_i$ is gray.

b. Either Sam doesn't own a donkey$_i$ or it$_i$ is gray.

There as well, it is an empirical question whether these sentences receive an existential reading (as in (7a) or (8a)), a universal reading (as in (7b) or (8b)) or both, or some other truth-conditions entirely. The goal of this work is to answer this empirical question.

(7) **Paraphrase of (6a):** *Sam owns at least one donkey and …*

a. … *at least one donkey owned by Sam is gray* (∃ reading)

b. … *every donkey owned by Sam is gray* (∀ reading)

(8) **Paraphrase of (6b):** *either Sam doesn't own a one or more donkeys or …*

a. … *at least one donkey owned by Sam is gray* (∃ reading)

b. … *every donkey owned by Sam is gray* (∀ reading)

We believe there are multiple theoretical benefits to answering the question. First, perhaps in part due to the literature's focus on donkey sentences, some theories rely on the presence of *every* to generate the relevant readings of (1) (Champollion et al., 2017; Chierchia, 1992; Heim, 1982, a.o.). By studying constructions that, on the surface, do not include any quantifier apart from the indefinite antecedent, we can decide which readings are attributable to the presence of a quantifier, and which ones are not.

Second, the theoretical literature contains contradicting predictions regarding the truth-conditions exhibited by bathroom sentences specifically[1]. Some theories expect bathroom sentences to receive universal readings, some to receive existential readings. By eliciting clear data, we hope to settle this debate.

This work presents three experimental studies investigating the truth-conditions that arise in cross-conjunction and bathroom sentences. To foreshadow, we conclude that there is evidence that bathroom sentences are ambiguous between an existential reading and a universal reading and that there is only evidence for an existential reading for cross-conjunction sentences. We argue that our results are challenging to all theoretical approaches.

The roadmap is as follows: in section 2, we present experiment I, a truth-value judgment task testing the truth-conditions of cross-conjunction sentences. Section 3 presents experiment II, which investigated the truth-conditions of bathroom sentences. The experiment III of section 4 investigates both types of sentences in tandem. It was aimed to rule out a potential interpretation of experiment II and to replicate the results from the previous two studies. Section 5 discusses the significance of our results with respect to existing theories: we conclude that, while our results fall in line with (a modified version of) Kanazawa (1994)'s generalization, they are not predicted by any extant theory.

2 Experiment I: pronoun across conjunction

In experiment I, we studied pronouns in conjunctive sentences like:

(9) There is a donkey$_i$ and it$_i$ is gray.

(10) **Readings:**
There is one or more donkeys and …

a. *…at least one donkey is gray.* (existential)

b. *…every donkey is gray.* (universal)

c. *…the unique donkey that there is is gray.* (uniqueness)

There exists a strong consensus, particularly in dynamic analyses (Groenendijk and Stokhof, 1990; Heim, 1982; Muskens, 1996), that such cross-conjunction pronouns yield existential readings, as in (10a). The main goal of this study is to confirm this conclusion in a controlled experimental setting with naïve participants. A secondary goal of this study is to test whether the sentence may also have other readings in addition to the existential reading. There are echoes to this in the theoretical literature. For instance, there are marginal claims (Chatain, 2018; van der Does, 1993) that universal readings in (10b) are possible. Taking inspiration from Kadmon (1990), we also seek to test whether speakers have detectable uniqueness intuitions and access a reading like (10c). Furthermore, we wonder whether such uniqueness intuitions stem from the use of the pronoun *per se*, or a uniqueness implicature coming from the singular indefinite (Sauerland, 2003; Spector, 2007), which is independently attested. To test this, we add, as a baseline, minimally different sentences that contain an indefinite but no pronoun. That way, we can compare between the amount of uniqueness readings accessed in sentences with and without pronouns.

2.1 Pre-registration

The study was pre-registered on the Open Science Framework (Foster and Deardorff, 2017) and is accessible at https://osf.io/n4vft.

2.2 Materials

2.2.1 Task

The task is a truth-value judgement task. Participants are presented with a picture and are asked to give a 7-point rating, on scale that ranges from *completely false* to *completely true*. Figure 1 illustrates a trial. Prior to the trials, the instructions (reproduced in (11)) explained to participants that they should rate "*how true the sentence feels to [them]*". To emphasize the non-normative nature of the task, we added that there was no correct

answer[2].

(11) For each trial, you will see several geometrical objects and a sentence in bold. Each time you will have to rate how true the sentence feels to you given the set of objects. There is no "correct answer". What we are interested in is you*[sic]* intuituve*[sic]* judgement.

Across all conditions, the picture presented shows 4 colored shapes displayed in a 2×2 grid (cf fig. (1)).

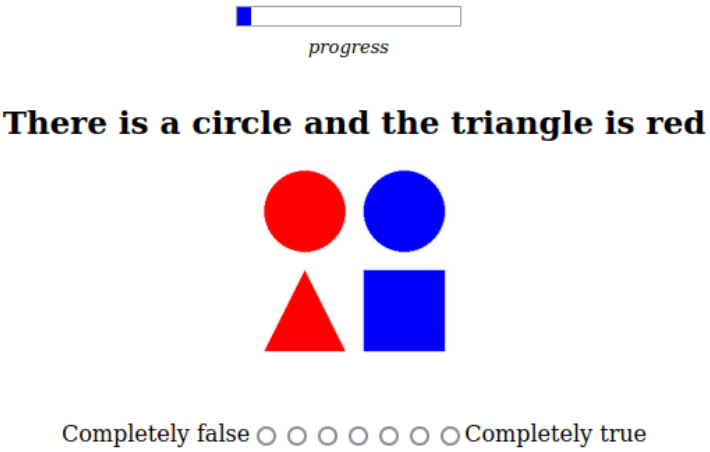

Figure 1: A typical trial in experiment I (condition NoPronoun-TrueWeak)

It may seem surprising to use scales to elicit truth/falsity judgments, instead of providing a binary choice between *true* and *false*. However, there is a good reason to think that scales are a more sensitive metric. First of all, scales allow participants to use mid-scale values to signal an ambiguity: a participant accessing a reading that is true of the depicted situation and one that is false may choose intermediate values to signal this ambiguity. There is prior evidence for this: (Marty et al., 2015) argue that, in cases of ambiguity where one reading is much preferred to another, scales can reveal a marginal

[2]Unfortunately, our instructions contained some typoes which were only caught after the experiments were run. These are, to our knowledge, the only typoes and we don't think they may affect the results in any way.

reading 2.check, where a binary choice may not. Second, Champollion et al. (2019) have claimed that anaphora may lead to undefined truth-conditions ; in their view, the different "*readings*" of a sentence correspond to different resolutions of the undefined truth-value. Scales can accommodate this theoretical possibility by letting participants express that a scenario is neither true nor false (but undefined). Third, and relatedly, it is known that providing more options than *true* and *false* may lead one of implicatures 3.cite Degen & co.. As stated in the introduction 4.write, we remain agnostic as to the underlying nature of the existential and universal intepretations (*bona fide* readings, different resolutions of a homogeneity gap, pragmatically strengthening). But given the evidence laid out before, we can see that, regardless of the exact nature of these interpretations, a sentence receiving both interpretations should yield intermediate or lower values when one of the interpretation is true and the other false. This should lead to detectable differences. In addition to all of this, there is a precedent for using scales in experimental studies on donkey anaphora from Denić and Sudo (2022). 5.check language used to talk about readings

2.2.2 Conditions

There are 9 conditions, which vary the type of sentence presented and the type of picture. All participants see all conditions (within-participant design). Among the 9 conditions, we have a "*pronoun*" group of 4 conditions and a "*no-pronoun*" group of 5 conditions. The two groups vary in the type of sentence presented. In conditions of the "*pronoun*" group, the sentence presented is of the form in (12), where SHAPE is randomly picked between either "*triangle*", "*circle*" or "*square*" and COLOR between with "*red*", "*green*" or "*blue*".

(12) **Pronoun conditions**
There is a SHAPE and it is COLOR.

The 5 conditions are described in table 1 (assuming SHAPE is triangle and COLOR is blue). The pictures are designed to tease apart the three readings we are interested in: the existential, the universal and the uniqueness reading. For instance, speakers accessing an existential reading would judge PRONOUN-EXISTENTIAL true and PRONOUN-SECOND-FALSE false, which would likely translate in a significantly higher mean rating for the former condition than for the latter.

(13) a. PRONOUN-FIRST-FALSE: the picture contains no triangle.

b. PRONOUN-SECOND-FALSE: the picture contains a unique triangle and that triangle is not blue.

	existential	universal	uniqueness
Pronoun-First-False	F	F	F
Pronoun-Second-False	F	F	F
Pronoun-Existential	T	F	F
Pronoun-Universal	T	T	F
Pronoun-Unique	T	T	T

Table 1: Readings true in each condition

c. Pronoun-Existential: the picture contains one blue triangle and one non-blue triangle and does not contain any other triangle.

d. Pronoun-Universal: the picture contains two blue triangles and does not contain any other triangle.

e. Pronoun-Unique: the picture contains one blue triangle and does not contain any other triangle.

As discussed in the previous section, conditions of the "*no-pronoun*" group served as a baseline for the pronoun conditions, to test any uniqueness implicature that may arise independently of the use of a pronoun. The sentence used in these conditions, given in (12), uses "*the* Shape$_2$" instead of *it* (where Shape$_2$ was always different from Shape$_1$).

(14) **No-pronoun conditions**
There is a Shape$_1$ and the Shape$_2$ is Color

To describe the conditions, we assume for concreteness that the sentence is *there is a triangle and the circle is blue*. In all pictures of the no-pronoun conditions, there is only one circle, thus the uniqueness presupposition carried by *the circle* is always met. There are four conditions:

(15) a. NoPronoun-First-False: the picture contains no triangle, a unique circle, which is blue and three other squares of any color.

b. NoPronoun-Both-False: the picture contains no triangle, a unique circle, which is not blue, and three other squares of any color.

c. NoPronoun-TrueWeak: the picture contains two triangles a unique circle which is blue and two other squares of any color.

d. NoPronoun-TrueStrong: the picture contains a unique triangle, a unique circle which is blue and three other square of any color.

There was 3 trials for each of the 9 conditions, amounting to 27 trials in total (no filler trials were used).

2.3 Participants

60 participants were recruited using the platform Prolific (Palan and Schitter, 2018). Through two question at the end, we excluded any participant who reported a form of color blindness or reported not being native speakers of English. As a first attention check, we excluded any participant who, on more than one trial, did not give one of the two lowest ratings to the NoPronoun-First-False and Pronoun-First-False conditions. We furthermore excluded participants who always answered with one of the two leftmost scale items for all trials. Neither condition was used as a basis for statistical comparison, so as to avoid spurious effects[3]. 4 participants ended up excluded by these criteria.

2.4 Results

Figure 2 represents the mean score given by participants to each condition. Figure 3 represents the distribution of ratings given by participants in each condition.

The statistical comparisons described below are obtained using the same methodology as in experiment I, which is described in section 2. We remind the reader that, for reasons explained in section 2, the tests do not correspond to the tests described in our pre-registration. The results of the latter are given in appendix 6.. But this seems inconsequential as the conclusions for this experiment are exactly identical whether we use the pre-registered tests or the ones described below.

Existential and universal readings. The sentence containing a pronoun was rated significantly lower in the Pro-2nd-False than in the Pro-∃ condition (two-sided paired t-test[4] ; t = 28.228, df = 108.55, p-value < $2.2e^{-16}$). The pictures in the Pro-∃ were

[3] There may still be a worry that participants' response to the Pronoun-First-False is correlated with their response to Pronoun-Second-False and that excluding participants giving too often a high score to the former condition might exclude participants who give a high score to the Pronoun-Second-False condition, creating an artificial difference between Pronoun-Second-False and Pronoun-Existential (one of our statistical comparisons). However, in post-hoc analyses, we find our results remain qualitatively the same if we don't perform any exclusions. So the four excluded participants are unlikely to be driving the difference we're finding.

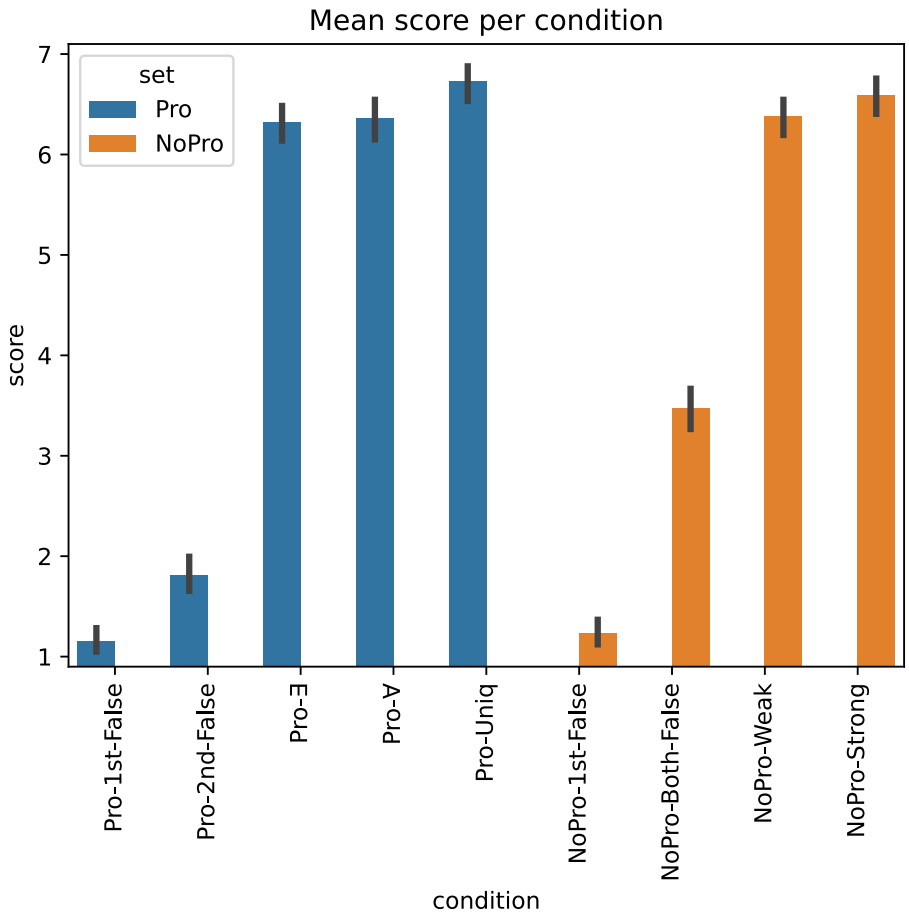

Figure 2: Mean score given by participants (1-7) to each condition. The conditions are divided in the PRO conditions where the sentence is of the form "*There is a triangle and it is blue*" and the NOPRO conditions where the sentences is of the form "*There is a triangle and the circle is blue*"

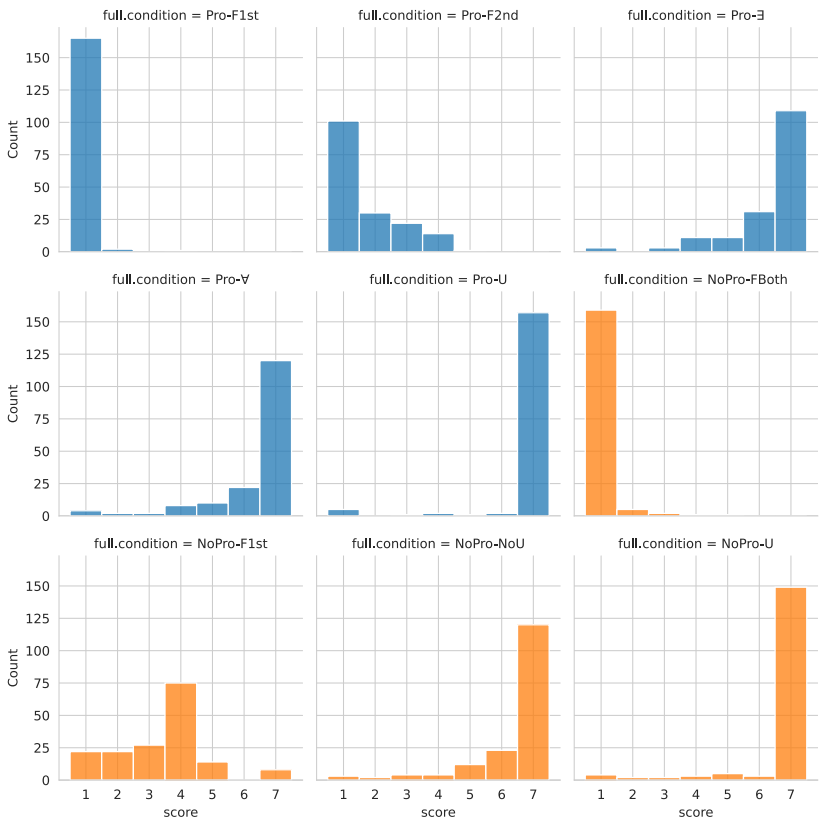

Figure 3: Answers given to each condition. Each bar represents the number of times a particular answer was selected

given on average 4.595 more points than the pictures in the PRO-2ND-FALSE condition (95% confidence interval: [4.27, 4.92]). On the contrary, the difference between the PRO-∃ and the PRO-∀ was not significant (two-sided paired *t*-test ; t = −0.24521, df = 105.09, p-value = 0.8068). The mean difference in ratings was only 0.047619 (95% confidence interval: [−0.432, 0.337]).

These results support the presence of an existential reading but do not provide evidence for the presence of a universal reading. One conclusion is that the universal reading does not indeed exist. Another *prima facie* possible explanation is that speakers do perceive both an existential and a universal interpretation but charitably interpret the sentence with the weaker existential interpretation in the PRONOUN-EXISTENTIAL where existential interpretation is true and the universal interpretation isn't. This alternative explanation faces challenges. First, as we argued in section 7., scales are known to be sensitive to ambiguity where one reading is much preferred to another (Marty et al., 2015). Second, as we'll see in experiment II (section 3) and III (section 4), we can reveal an ambiguity in disjunction configurations with the same set-up used here. If a "charitable interpretation" strategy were available, it is a mystery why it would only be available in conjunctions but not in disjunctions.

We may now turn to the question of whether there is an effect of uniqueness and whether this effect is mediated by the presence of a pronoun.

Uniqueness effects. The score given to the PRO-∀ was rated on average 0.38 lower than in the PRO-∀ condition but the difference between the PRO-∀ and the PRO-UNIQ conditions did not pass the 0.05 significance threshold after correcting for multiple comparisons (two-sided paired *t*-test ; t = 2.0146, df = 101.85, p-value = 0.1397).

Finally, we wanted to check whether the presence of the pronoun mediated any effect of uniqueness. To do so, we performed a 2-way within subjects ANOVA, with group (pronoun vs no pronoun) as the first independent variable and condition type (unique vs not unique) as the second independent variable. We chose the PRO-∀ to stand for the non-unique condition in the pronoun group (with NOPRO-NOUNIQ being its counterpart in the no-pronoun condition). The interaction was not significant (F(df interaction, df within) = 0.953, p = 0.666).

[4] We run the *t*-tests and 2-way ANOVA as follows. First, for each and each condition, we average the score given to the three trials of that condition by that subject. The paired *t*-test wouldn't be applied otherwise. Then, we subset our data to the two conditions (*t*-test) or two pairs of conditions (ANOVA) we wish to compare. We run the test on this averaged subsetted dataset.

2.5 Discussion

Overall, our results strongly confirm the presence of an existential reading of cross-conjunction anaphora. We do not find evidence of a universal reading. This could either mean that this universal reading simply does not exist or that it requires special licensing conditions which were not met in our experiment set-up.

Likewise, the effect of uniqueness could not be detected. The differences in score between the Pro-∀ condition where the sentence "*There is a triangle and it is blue*" is matched with a picture with two blue triangles, and the Pro-Uniq where it is matched with a picture with a single triangle, which is blue, did trend in the expected direction: the score given to the former was on average lower than in the unique condition. However, this difference was not significant. Furthermore, there was no significant interaction between use of a pronoun and uniqueness. We thus do not find evidence that the pronoun, by itself, imposes uniqueness restrictions on top of any uniqueness implicatures that the indefinite itself may yield.

These results overall comfort the conclusions of the literature, including but not limited to the dynamic literature, which predict existential truth-conditions for pronoun across conjunctions. In the sequel, we investigate the truth-conditions of pronouns across *disjunctions* (so-called *bathroom sentences*), which are more controversial.

3 Experiment II: pronoun across disjunction

In this second experiment, we are interested in the truth-conditions of sentences like (16), so-called *bathroom sentences*, where the pronoun and its indefinite antecedent are split across a disjunction. The felicity of such sentences raises a challenge to classical dynamic treatments (Groenendijk and Stokhof, 1990, 1991; Heim, 1982) and these sentences have been heavily discussed in that connection.

(16) Either there isn't a circle or it is blue.

a. **Existential reading:**
If there is at least one circle, one circle is blue.

b. **Universal reading:**
If there is at least one circle, every circle is blue.

c. **Uniqueness reading:**
If there is at least one circle, there is just one and it is blue.

Many theories have been proposed to explain why bathroom sentences sentences are felicitous (Elliott, 2020; Krahmer and Muskens, 1995; Mandelkern, 2020). Somewhat less discussed is the question of which truth-conditions bathroom sentences should receive. As with conjunction, three readings may *a priori* be expected: an existential (16a), a universal (16b) and a uniqueness reading (16c).

Our goal was to test which reading or readings among these three were actually accessed by naïve participants. While we could have tested sentences of the form in (16), we worried about possible ignorance inferences of disjunction. In a natural setting, a cooperative speaker can only utter a disjunction like (16) if they do not know which disjunct is true, in particular whether there was a circle or not. Since pictures fully specify all the relevant information, participants might reject the sentence, simply on the basis that a speaker would never have uttered the sentence, if they are cooperative and see the same picture that the participant sees.

To remedy this potential confound, we tested the slightly more complicated sentences in (17), embedding the disjunction under the quantifier *in every row*. When embedded under quantifiers[5], disjunctions do not yield *ignorance inferences* ; they could be uttered by a fully informed speaker. Instead, they give rise to *distributive implicature*[6], as presented in (17), which will be met in all the pictures we construct.

(17) In every row, either there isn't a circle or it is blue.

a. **Existential reading:**
In every row with one circle or more, at least one circle is blue.

b. **Universal reading:**
In every row with one circle or more, every circle is blue.

c. **Uniqueness reading:**
In every row with one circle or more, there is just one circle and it is blue.

(18) a. $\rightsquigarrow$ *in some row, there isn't a circle*

b. $\rightsquigarrow$ *in some row, there is a blue circle*

[5]Ignorance inferences can still arise in this context if the domain of quantifier is smaller than the number of disjuncts (Denic, 2020). But such is not the case in our examples.

[6]Interactions between implicatures of disjunction and pronouns has, to our knowledge, only been discussed in Elliott and Sudo (2023), in the case of Free Choice specifically. We don't know whether the intuitions we informally report here for distributive implicatures are predicted by any theory.

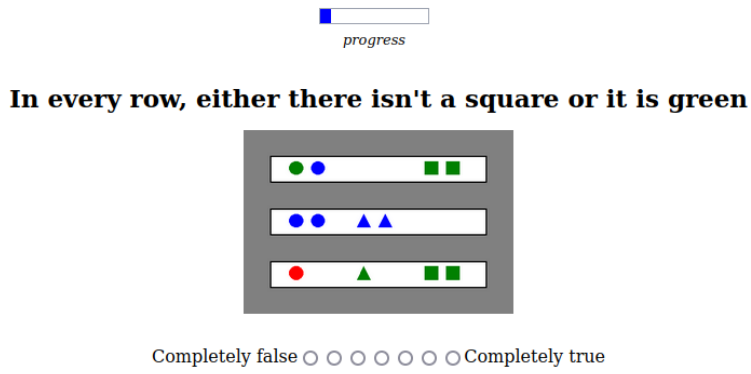

Figure 4: A typical trial in experiment II

A final remark is that as with conjunctions, we may ask whether the presence of a uniqueness reading could be attributed to an implicature of uniqueness, arising independently from the singular indefinite *a circle*. However, that proposal is harder to formulate in the bathroom case than in the conjunction case, because the indefinite *a circle* here is embedded in a downward-entailing environment. In these environments, no uniqueness implicature may arise[7]. For this reason and to keep the load of the experiment light on participants, we chose not to include no-pronoun baseline conditions, as we did in the experiment I of section 2.

3.1 Pre-registration

This study was pre-registered on the Open Science Framework (Foster and Deardorff, 2017) and is accessible at https://osf.io/8jzs4.

3.2 Materials

The task was exactly the same as in experiment I from section 2, but for the pictures and sentences used.

The pictures now included rows of colored shapes, with varying number of shapes per row, as figure 4 illustrates. To make the picture easier to parse, each type of shape was placed in the same position in each row and each trial: circles were on the left, triangles in the middle and squares on the right.

[7] It may be that a weaker implicature is derived, namely *if there is a circle, there is only one*. We acknowledge this might be a possibility but we don't know of any theory that would predict such an inference.

The sentences presented were always of the form in (19).

(19) In every row, either there isn't a SHAPE or it is COLOR

There were 5 conditions, 3 trials per condition, amounting to 15 trials in total. The conditions only varied the type of picture presented. There were two false baselines, DISJ-F1ROW and DISJ-F2ROWS, depending on the number of rows for which the disjunction was false. DISJ-F2ROWS was used for our exclusion criteria[8] and DISJ-F1ROW was used as basis for statistical comparisons. In DISJ-F1ROW, one row made disjunction false and two rows made it true ; in DISJ-F2ROWS, two rows made it false and one made it true. Taking (17) as a reference sentence, false rows contained a single non-blue circle and other non-circle shapes. Regardless of the reading accessed for disjunction, we expected this to make the conjunction false. In true rows, there was no circle at all.

The three target conditions were DISJ-∃, DISJ-∀ and DISJ-U. In each target condition, two rows made the disjunction true according to one reading but not any stronger reading, one row made it true by not having any circles. So for instance, two rows in the DISJ-∃ picture made the existential reading true and not the universal one (by having one red circle and one blue circle) and one row contained no circle at all. We summarize the different conditions in the list below:

- DISJ-F2ROWS: 2 rows with just 1 non-blue circle, 1 row with no circle false under all
- DISJ-F1ROW: 1 row with just 1 non-blue circle, 2 rows with no circle
- DISJ-∃: 2 rows with 1 blue circle, 1 non-blue circle, other rows no circle
- DISJ-∀: 2 rows with 2 blue circles, other rows no circle
- DISJ-U: 2 rows with 1 blue circle, other rows no circle

In table 2, we list how the different conditions would be judged by participants depending on the readings they access.

[8] As for Experiment I (cf fn. 3), there may be a worry that the answers to DISJ-F2ROWS and DISJ-F1ROW likely are correlated so that by excluding participants who give a high score to DISJ-F2ROWS, we're excluding participants who give a high score to DISJ-F1ROW, with the potential to create an artificial difference between DISJ-F1ROW and the DISJ-∃ (one of the comparisons to be run). However, in post-hoc analyses, we find our results remain qualitatively the same even if we run our analyses with all participants, making it unlikely that the effects observed are the result of our exclusion procedure.

	existential reading	universal reading	uniqueness readings
Disj-F2Rows	F	F	F
Disj-F1Row	F	F	F
Disj-∃	T	F	F
Disj-∀	T	T	F
Disj-U	T	T	T

Table 2: Predicted truth-values for each condition, depending on the reading of the sentence

3.3 Participants

80 participants were recruited using the platform Prolific (Palan and Schitter, 2018). Through two questions at the end, we excluded any participant who reported a form of color blindness or reported not being native speakers of English. First, we excluded any participant who, on more than one trial, did not give one of the two lowest ratings to the Disj-F2Rows. Second, we excluded participants who always answered with one of the two leftmost scale items for all trials. 13 participants ended up excluded by these criteria.

3.4 Results

All statistical tests presented here were pre-registered, unless explicitly noted. The Holm-Bonferroni correction for multiple testing was used ; the p-values reported below are the corrected p-values.

Figure 5 represents the mean score given by participants to each condition.

The Disj-∃ condition was rated significantly higher than the Disj-F1Row condition (two-sided paired t-test[9] ; t = 4.2049, df = 126.45, p-value = $9.8e^{-5}$) but significantly lower than the Disj-∀ (two-sided paired t-test ; t = -6.7844, df = 125.5, p-value = $1.2e^{-9}$). The difference between the Disj-∀ and the Disj-U conditions was not significant (two-sided paired t-test ; t = 0.1762, df = 130.95, p-value = 0.8604).

From this, we conclude that the disjunction is truly ambiguous between an existential reading and a universal reading. In certain circumstances, speakers access a universal reading leading them to judge the sentence false and, in others, they access an existential reading and judge it true. As a result, the mean score for the Disj-∃ is intermediate between Disj-F1Row and Disj-∀ reading. To corroborate that interpretation, we may observe that the responses to the Disj-∃ seem to follow a bi-modal distribution (cf

[9]The methodology to run these test is as described in fn. 4

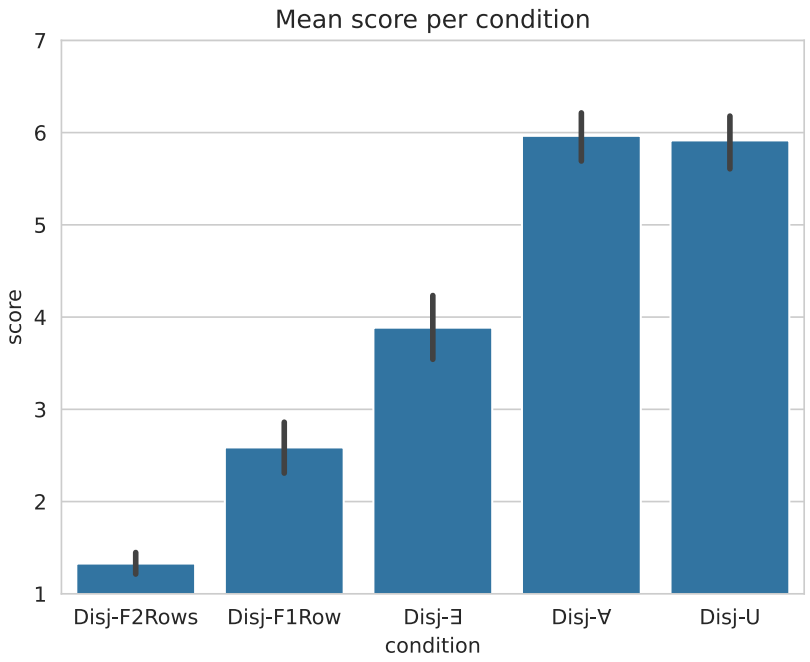

Figure 5: Mean score given by participants (1-7) to each condition.

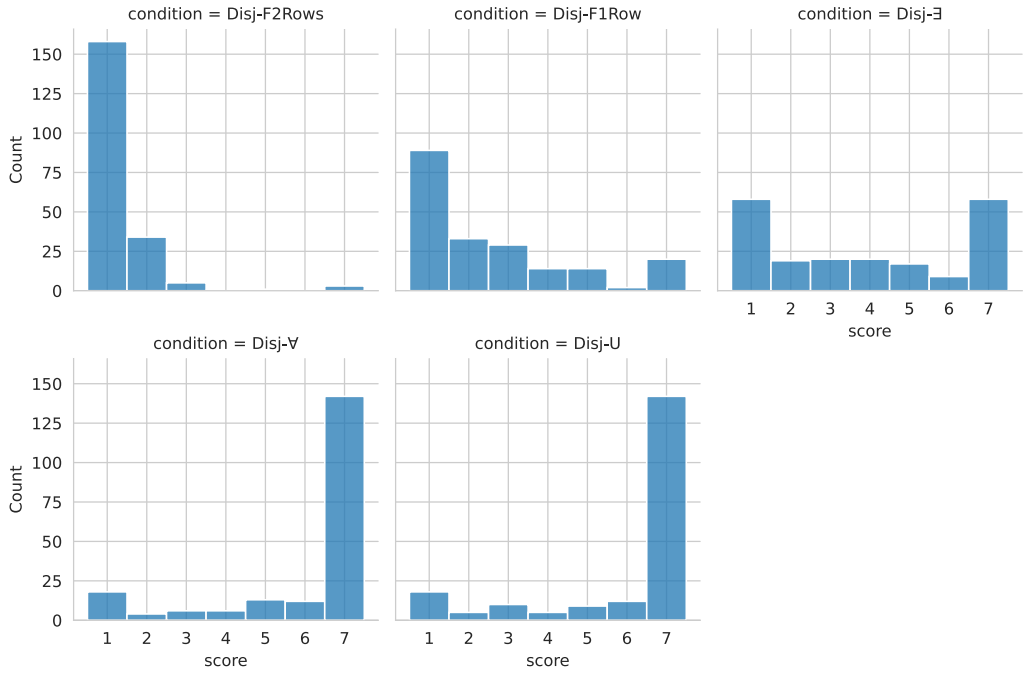

Figure 6: Answers given to each condition. Each bar represents the number of times a particular answer was selected

fig. 6), with the lowest score and the highest score being the two most selected responses.

3.5 Discussion

The results of this experiment confirm the presence of both an existential and a universal reading for bathroom sentences. If true, these results are interesting because they are challenging for all existing theories, as we'll discuss in section 5. In the next experiment, we seek an explicit comparison between the conjunction case and the disjunction case and to rule out a possible interpretation of experiment II.

4 Experiment III: pronouns across conjunction and disjunction

This study compares the readings of pronouns across conjunctions (cross-conjunction anaphora) and disjunctions (*bathroom sentences*). Our first goal was to confirm the dis-

crepancy suggested by the previous two experiments: while we could only evidence an existential reading in conjunctions (experiment I in section 2), we found evidence for both an existential and a universal reading in disjunctions (experiment II in section 3). We now wanted to replicate these results within participants, by presenting both types of sentences within one and the same experiment.

Our second goal was to rule out a possible interpretation of experiment II. Because the bathroom sentences presented in experiment II contained the quantifier *in every row*, that universal quantifier might be argued to be the source of the universal reading, rather than it being a property of disjunction *per se*. This alternative explanation would have theoretically precedents: chapter 2 of Heim (1982) essentially proposes that universal quantifiers are unrestricted binders and may bind both indefinites and pronouns. The idea comes from Lewis (1975) and appears in many other works (Kamp et al., 2011; Schubert and Pelletier, 1989, a.o.). Under this view, bathroom sentences may be represented by (20a) ; with this parse, they would receive the truth-conditions in (20b), which are equivalent to the universal reading.

(20) a. In every row^x, either there isn't a circle$_x$ or it$_x$ is blue.

b. $\forall y, \forall x, \mathsf{row}(y) \rightarrow \neg(\mathsf{circle}(x) \wedge \mathsf{in}(y)(x)) \vee \mathsf{blue}(x)$

To test this alternative interpretation, we tested the readings of unembedded disjunctions directly. If the universal quantifier *in every row* were critical to the generation of the universal reading, unembedded disjunction should lack a universal reading, resulting in higher ratings to pictures that make the existential reading true and the universal reading false.

As explained in section 3, we initially shun away from unembedded disjunctions, as potential ignorance inferences may lead participants to judge such unembedded disjunctions to be infelicitous. However, we reasoned that any such degradation, if it were found, should affect all true conditions of the disjunctive sentences indiscriminately and, in particular, would not change the interpretation of any comparison with the conjunction cases.

4.1 Pre-registration

This study was pre-registered on the Open Science Framework (Foster and Deardorff, 2017) and is accessible at https://osf.io/edmgx.

4.2 Materials

The task was the same picture rating task as in experiment I and experiment II. The pictures were like those used in experiment I, displays of four colored shapes (cf fig. 1). Each trial either belonged to the Conj group or Disj group. The sentence used for the Conj trial was of the form in (21a), while the sentence used for the Disj trial was of the form in (21b).

(21) a. There is a Shape and it is Color.

b. Either there isn't a Shape or it is Color.

In trials of the Conj group, the picture presented was one of the following 5 picture types (illustrated for *there is a triangle and it is blue*) in (22). The picture types described in (22a-b) are two false baselines, making false either the the first conjunct or the second one (hence their names) ; the other 3 are the target conditions. Table 3 lists the predicted truth-values for each of these conditions based on the reading accessed.

(22) Conj conditions

a. Conj-F1st: 4 non-triangle shapes of a random color.

b. Conj-F2nd: exactly one triangle, the triangle is not blue, 3 non-triangle shapes.

c. Conj-∃: exactly two triangles, one blue and one non-blue, 2 other non-triangle shapes.

d. Conj-∀: exactly two triangles, both blue, 2 other non-triangle shapes.

e. Conj-U: exactly one triangle, the triangle is blue, 3 other non-triangle shapes.

For the Disj group, there were 5 picture types. They are given in (23) (for the sentence *either there isn't a triangle or it is blue*). Table 4 lists the truth-values predicted for each of these conditions based on the reading accessed.

(23) a. Dist-F: exactly one triangle, the triangle is not blue, 3 non-triangle shapes.

b. Disj-∃: exactly two triangles, one blue and one non-blue, 2 other non-triangle shapes.

c. Disj-∀: exactly two triangles, both blue, 2 other non-triangle shapes.

d. Disj-U: exactly one triangle, the triangle is blue, 3 other non-triangle shapes.

e. Dist-T1st: no triangle, 4 non-triangle shapes.

	existential reading	universal reading
Conj-F1st	F	F
Conj-F2nd	F	F
Conj-∃	F	T
Conj-∀	T	T
Conj-U	T	T

Table 3: Predicted truth-values for each Conj conditions, depending on the reading of the sentence

	existential reading	universal reading
Dist-F	F	F
Disj-∃	F	T
Disj-∀	T	T
Disj-U	T	T
Dist-T1st	T	T

Table 4: Predicted truth-values for each Disj conditions, depending on the reading of the sentence

There were 3 trials for each condition, amounting to $3 \times (5 + 5) = 30$ trials. The experiment was split in two blocks: one block only contained Conj trials, one block only contained Disj trials. The two blocks were separated by a screen indicated "*We now move on to a different type of sentence*". This was done in an effort to minimize workload, as all sentences within one block were of the same form. The order of the Conj and Disj block was randomized across participants.

4.3 Participants

130 participants[10] were recruited using the platform Prolific (Foster and Deardorff, 2017). Through two question at the end, we excluded any participant who reported a form of color blindness or reported not being native speakers of English. As a first attention check, we excluded participants who scored more than 2 on at least two of the Conj-F1stcondition trials. As a second attention check, we excluded any participants who scored less than 6 on at least two of the Dist-T1st condition trials.

[10] More participants recruited here than in the previous two experiments. This number was calculated by a power analysis, using the results of the previous two experiments as a basis for the calculation. Since the difference between the uniqueness and the universal conditions in experiment I was very small (and

With this criterion, we ended up excluding 61 participants. This represents ~47% of our participants. It seems that 59 of the excluded participants were excluded on the basis of the second attention check ; in other words, roughly half of the participants gave a low (false-like) rating to a sentence like (24), when the picture contained no circles at all (i.e. the first disjunct is true).

(24) Either there isn't a circle or it is blue.

The large number of exclusions is worrying. It warrants a thorough discussion, which we conduct in the discussion section 4.5. To foreshadow, we argue that even with this high exclusion rate, the result still allows to conclude to the presence of a universal reading in bathroom sentences, independently of the presence of a quantifier like *in every row* (as used in experiment II), which answers the question we sought to answer. We set aside this discussion for the time being and present our statistical comparisons.

4.4 Results

Figure 7 represents the mean score given by the non-excluded participants for each condition. All statistical tests presented here were pre-registered, unless explicitly noted. Here, every comparison was obtained by the following method: we fitted Cumulative Link Models (Agresti, 2012) to the data[11] including participant as the random effect and then performed likelihood ratio tests in which the deviance of the models containing the main or interaction effect of interest was compared to another model without that effect. We report the χ^2 and p-values obtained in doing such comparisons. The Holm-Bonferroni correction method for multiple comparisons was used ; the p-values reported below are therefore corrected p-values.

Overall, we mostly replicated the results of experiment I on conjunctions and experiment II on disjunctions in the corresponding block. For the Conj block, we found that speakers gave significantly higher ratings to the Conj-∃ than to the Conj-F2 ($\chi^2(\text{df} = 1) = 536.37, p \leq 2.2e^{-16}$), but no significant difference were found between the Conj-∃ and the Conj-∀ ($\chi^2(\text{df} = 1) = 0.3841, p = 0.5354$). The difference between the Conj-∀ and the Conj-U was significant ($\chi^2(\text{df} = 1) = 47.863, p = 2.286e^{-11}$). This is an effect that we did not find in experiment I.

indeed not significant), it was found that any effect of uniqueness would require twice as many participants to evidence.

[11] Unlike the methodology reported in the previous two experiments, these tests correspond to the pre-registered tests.

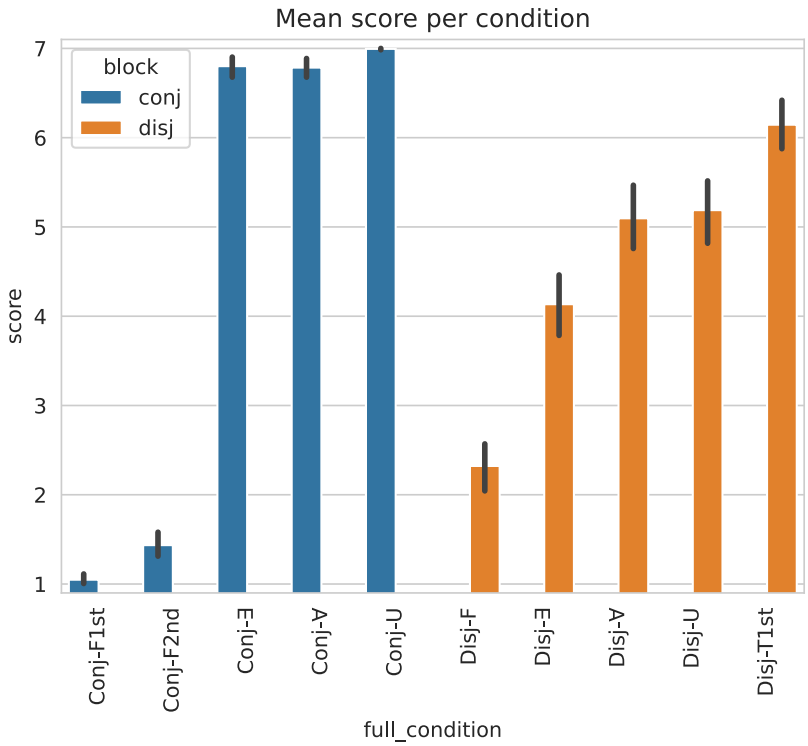

Figure 7: Mean score given by participants (1-7) to each condition.

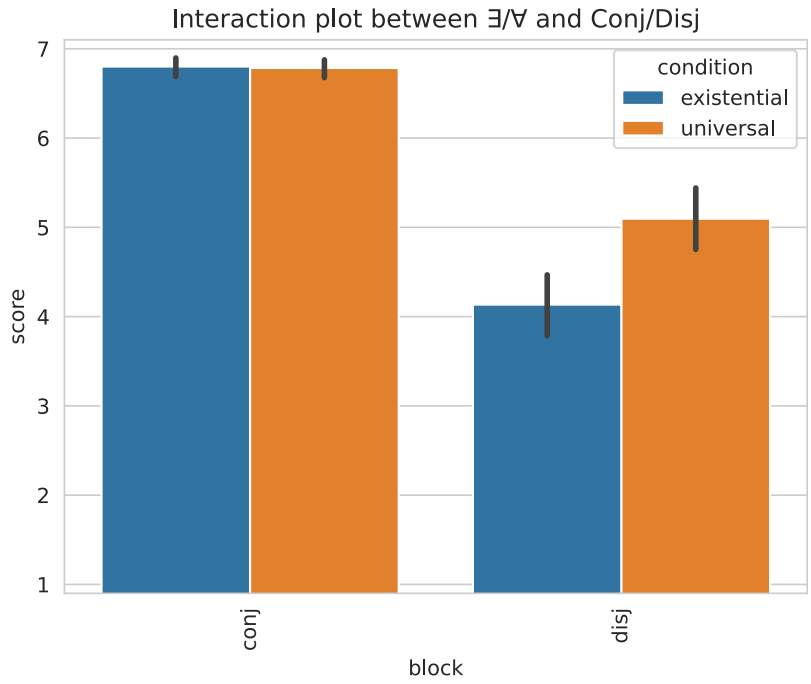

Figure 8: Interaction plot showing comparing the difference between the ∃ and the ∀ condition in the Conj block and the Disj block.

For the Disj block, we found that speakers gave significantly higher ratings to the Disj-∃ than to the Disj-False ($\chi^2(\text{df} = 1) = 69.103, p \leq 2.2e^{-16}$), and that they gave significantly higher ratings to the Disj-∀ than to the Disj-∃ ($\chi^2(\text{df} = 1) = 29.98, p = 1.746e^{-7}$). No significant difference were found between the Disj-∀ and theDisj-U ($\chi^2(\text{df} = 1) = 1.2348, p = 0.267$).

Finally, we found a significant interaction between picture (∃ vs. ∀) and block (Disj vs Conj) ($\chi^2(\text{df} = 1) = 9.0431, p = 0.00791$). The interaction plot is given in fig. 8.

4.5 Discussion

Overall, this experiment seems to corroborate the results of the previous two. In conjunctive sentences, only existential readings can be evidenced. In disjunctive sentences, both an existential reading and a universal reading are detected.

But there is a caveat. Roughly half of the participants were excluded because they judged the disjunctive sentence false in the control condition when there were no cir-

cles (a normatively incorrect behavior). This is an indication that participants did not complete the task as intended. If they did not complete the task as intended, can we be sure that the differences observed between the target conditions truly reflect readings of the sentence? We want to spell this worry out in details ; to foreshadow, we argue that, all things considered, differences between the target conditions cannot solely be attributed to the unexplained behavior we observed in control condition.

First, we state that we don't have an explanation for why participants reject the disjunctive sentence in the Dist-T1st condition. There are many options: (a) participants could be rejecting the sentence because of unmet ignorance implicatures, as we discussed in the introduction of this section, (b) they could be misread the sentence as *there is a circle and it is blue* (a misparse), (c) they could have adopted a inaccurate but fast verification strategy (*give high mark if green circle visible*), or yet other hypotheses. While there may be reasons to choose one hypothesis over another[12], we want to reason generally: we call UB this unexplained behavior. Every participant has a certain probability p of exhibiting UB and they may exhibit it on both control and target trials. By definition, we can say participants exhibiting UB give low ratings to the Dist-T1st ; however, short of an explanation on what UB is, we don't know *a priori* what ratings participants which exhibit UB on target trials will give. If, for instance, UB leads to give low ratings to the Disj-F2Rows and high ratings to the Disj-∃, this may result in a significant difference between these two conditions, but this difference could not be attributed to the presence of a universal reading, because it is not known what UB is.

Our exclusion criterion does not guarantee that all participants remaining after exclusion don't exhibit UB. It simply guarantees that they did not exhibit UB more than once on a control trial. This makes it possible to spell out the worry somewhat more clearly: it could well be that the difference we observed between the Disj-F2Rows and the Disj-∃ conditions, is entirely attributable to UB.

While our exclusion criteria don't rule out non-excluded participants from exhibiting UB in target trials, it is reasonable to say this: *on average,* the probability p of exhibiting UB is higher in excluded participants than in included participants. This follows if, as is likely, participants don't all have the same probability of exhibiting UB. In that case, participants who are more liable to UB are also more likely to exhibit UB on control trials and thus more likely to be excluded.

This observation means that we can get an insight on what responses UB triggers by comparing the results of excluded participants to the results of non-excluded par-

[12] There are good reasons to think that (a) is not the case. If participants did reject the sentence on the basis of unmet ignorance implicatures, we'd expect a similar rejection in all conditions that the participant deems true. But we find that rejected participants, like included participants, give high ratings to the Disj-U condition, which is true under all readings the participants may get.

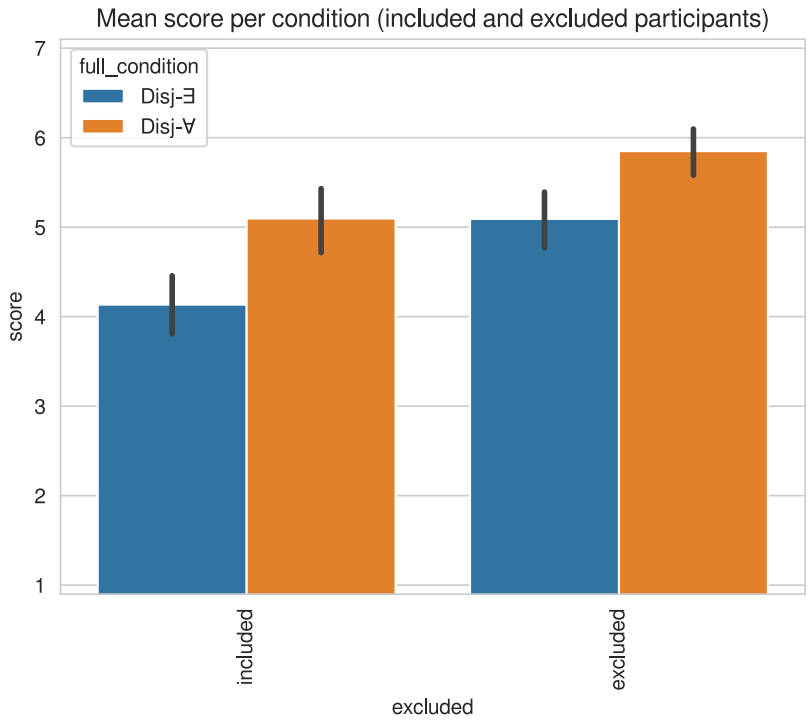

Figure 9: Interaction plot showing comparing the difference between the Disj-∃ and the Disj-∀ conditions across included and excluded participants.

ticipants. More specifically, if the differences observed between the target conditions is wider among excluded participants than they are among included participants, then we can conclude that UB might indeed be responsible for these differences. If, on the other hand, the differences are narrower in excluded participants, then it would suggest that our results are not attributable to UB.

Fig. 9 suggests the latter is the case for the difference between the Disj-∃ and the Disj-∀ conditions: the difference between the two conditions is actually narrower in the excluded participants than in the included participants. We can therefore conclude that at least the universal reading is not due to UB and thus that there is a universal reading in the absence of a universal quantifier.

Could it be that it is the existential reading that is spurious and due to UB? Experiment II already concluded to the presence of such readings (and furthermore, it is not possible to derive them relying on the quantification of the universal quantifier *in every row*). In this experiment, we can know that speakers did not access UB. In experiment

II, every condition included rows with no circles. UB, by definition, lead participants to judge such cases false. As a result, the sentence (25) would have come out false regardless of the condition and so it is unlikely that participants exhibited UB in this experiment[13]

(25) In every row, either there isn't a circle or it is blue.

In summary, we argue that while the presence of UB is real and troublesome, it does not invalidate our conclusion that both the existential and the universal readings of cross-disjunction anaphora are available in disjunctive sentences. Furthermore, we can conclude that the universal reading does not require the presence of a universal quantifier like *in every row*, showing it's a genuine reading of pronouns in disjunctions.

5 General Discussion

The conclusion of experiments I, II and III is that cross-conjunction anaphora is unambiguously read existentially, while bathroom anaphora is ambiguous between an existential and a universal reading. Is this pattern predicted?

We discuss three classes of theories: theories that predict that bathroom sentences are not possible with the given co-indexation (Groenendijk and Stokhof, 1991), theories that predict existential readings for bathroom sentences (Elliott, 2020), theories that predict universal readings for bathroom sentences (Krahmer and Muskens, 1995). We show a dilemma: while Elliott (2020) and Krahmer and Muskens (1995) may be amended to predict our results, this amendment comes at the cost of making ill-justified stipulations, ultimately undermining Elliott (2020)'s goal of achieving a "*principled*" explanation. On the other hand, adapting Elliott (2020)'s principled account runs the risk of predicting cross-conjunction anaphora to be ambiguous, contrary to our results.

5.1 Theories that don't predict bathroom sentences to be possible

Let us illustrate in more technical details the predictions of the different theories. We start with a theory that doesn't predict the possibility of bathroom sentences, the Dynamic Predicate Logic of (Groenendijk and Stokhof, 1991). In (our formulation of) this standard dynamic theory, a clause denotes a functions from input assignment to sets of output assignments (type ggt). A sentence is "*true*" if its denotation takes the input assignment given by the context to a non-empty set of assignments. A sentence like *there*

[13] Why didn't participants exhibit UB in experiment II? We speculate that the relative hardness of the task in experiment II might be responsible. compared to experiment I, made participants more attentive. If UB is due to a lack of attention, as we submit, this would explain this

is a circle$_i$ will, for instance, map an input assignment g to any assignment g' differing from g only in that it assigns i to a circle (formal definition in (26a)). Since this set is empty just in case there is no circle, it follows that the sentence will be true just in case there is a circle.

In DPL, negation checks what its prejacent would map the input assignment to: if the prejacent would map the input assignment to an empty set (i.e. the prejacent is false), then negation leaves the input assignment as is (which means the sentence is true) ; if, on the other hand, the prejacent can update the input assignment to a non-empty set (i.e. the prejacent is true), then negation returns an empty set (i.e. falsity). An important fact is that negation has the property of being "*externally static*": a constituent of the form "*not p*", no matter what *p* is, will never alter the assignment function in a way that may make new pronouns available: it either leaves the input assignment as is or maps it to the empty set.

(26) a. $[\![\text{there is a circle}_i]\!] = \lambda g.\lambda g'.\ \exists x, \text{circle}(x) \wedge g' = g_{[i \to x]}$

b. $[\![\text{it}_i \text{ is blue}]\!] = \lambda g.\lambda g'.\ g = g' \wedge \text{blue}(g_i)$

c. $[\![\text{not}]\!]\,(p_{ggt}) = \lambda g.\lambda g'.\ g = g' \wedge \neg\exists h, p(g)(h)$

d. $[\![\text{and}]\!]\,(p_{ggt})(q_{ggt}) = \lambda g.\lambda g'.\ \exists h, p(g)(h) \wedge q(h)(g')$

e. $[\![\text{or}]\!]\,(p_{ggt})(q_{ggt}) = \lambda g.\lambda g'.\ g = g' \wedge [\exists h, p(g)(h)] \wedge [\exists h, q(g)(h)]$

This definition of negation has the desirable consequence of ruling out impossible anaphoric link across negation, as in (27a). But it has the unfortunate consequence of ruling out bathroom sentences, where reference across negation is possible.

(27) a. # I don't have a plus one$_i$. They$_i$'re not on the guest list.

b. Either there isn't a bathroom$_i$ or it$_i$ is upstairs.

This is a well-known problem of DPL and similar theories. Because such theories don't predict bathroom sentences to be felicitous in the first place, they are trivially incompatible with our experimental results.

5.2 Universal reading of bathroom sentences (Krahmer and Muskens, 1995)

Some theories do solve the challenge of bathroom sentences and make concrete predictions regarding its truth-conditions. Krahmer and Muskens (1995), our first example,

propose to enrich the semantics by assuming that sentences may denote a pair of two updates[14], rather than just one as in DPL. In this system, a proposition denotes a pair of a positive update (which we write $[\![\cdot]\!]^+$) and a negative update (which we write $[\![\cdot]\!]^-$). For current purposes, we submit that a sentence is considered true if it can be updated by the positive update. With this, (26a) is enriched into (28): its positive update corresponds to the update of *there is a circle* as seen in DPL above and its negative update corresponds to the update of *there isn't a circle* in DPL.

(28) $[\![\text{there is a circle}_i]\!]^+ = \lambda g.\lambda g'.\ \exists x, \mathsf{circle}(x) \wedge g' = g_{[i \to x]}$
≈ *if there is a circle, add it at index i ; fail otherwise*
$[\![\text{there is a circle}_i]\!]^- = \lambda g.\lambda g'.\ g = g' \wedge \neg\exists x, x \in \text{circle}$
≈ *if there isn't a circle, don't update ; fail otherwise*

The use of positive and negative extensions allows for the simple definition of negation in (29) (as a so-called *flip-flop operator*). With this definition of negation, the positive update of the prejacent of negation S, and the discourse it may introduce, are still present in the negative update $[\![\cdot]\!]^-$. In principle, the discourse referents may be retrieved later in the composition. Another important consequence of this definition of negation is that double negations can be eliminated: *not not S* has the same positive and negative updates as S.

(29) $[\![\text{not S}]\!]^+ = [\![\text{S}]\!]^-$
$[\![\text{not S}]\!]^- = [\![\text{S}]\!]^+$

The semantics of the connectives also needs to be adapted to positive and negative updates. Krahmer and Muskens (1995)'s semantics of *or* and *and* is given in (30), focusing on the positive updates for ease of exposition. In a nutshell, *S or S′* is treated as having the same meaning as *If not S, then S′*.

(30) a. $[\![S \text{ and } S']\!]^+ = \lambda g.\lambda g'.\ \exists h, [\![S]\!]^+(g)(h) \wedge [\![S']\!]^+(h)(g')$
b. $[\![S \text{ or } S']\!]^+ = \lambda g.\lambda g'.\ g = g' \wedge \forall h, [\![S]\!]^-(g)(h) \rightarrow \exists h'\, [\![S]\!]^+(h)(h')$

We can now see that bathroom sentences are derived. Applying to the definition of *or* in (30b) to the meaning of the individual disjuncts in (31d) and (31b), we get the positive update in (31e). In plain language, (31a) will be true (i.e. its update won't result in the

[14] They don't couch their formalism using the kind of direct semantics we used here. Theirs is a form of DRT (Muskens, 1996). In order to avoid multiplying frameworks, especially since our point doesn't hinge on it, we translate it to the semantics we used here to describe DPL.

empty set) just in case every update of the context with the negative update of *there isn't a circle* can be followed by an update of the context with the positive update of *it is green*. Because of the definition of *not*, the negative update of *there isn't a circle*$_i$ is the positive update of *there is a circle*$_i$. As we saw above, this update introduces a discourse referent at index i corresponding to a circle. Overall, the sentence requires that every such update can be followed by a positive update of *it is green*.

(31) a. Either there isn't a circle or it is green.

b. $[\![\text{it}_\text{i} \text{ is green}]\!]^+ = \lambda g.\lambda g'.\ g = g' \wedge \text{green}(g_i)$

c. $[\![\text{there isn't a circle}_\text{i}]\!]^+ = \lambda g.\lambda g'.\ g = g' \wedge \neg\exists x, x \in \text{circle}$

d. $[\![\text{there isn't a circle}_\text{i}]\!]^- = \lambda g.\lambda g'.\ \exists x, \text{circle}(x) \wedge g' = g_{[i \to x]}$

e. $[\![\text{Either there isn't a circle or it is green}]\!]^+$
$= \lambda g.\lambda g'.g = g' \wedge \forall h, (\exists x \in \text{circle}, h = g_{[i \to x]}) \to (\exists h', h = h' \wedge \text{green}(h_i))$
$= \lambda g.\lambda g'.g = g' \wedge \forall x, x \in \text{circle} \to \text{green}(x)$

The prediction made is that the sentence is true just in case every circle is green (considered true when there is no circle). These are the universal truth-conditions. As presented, Krahmer and Muskens (1995) thus don't predict that bathroom sentences may have an existential reading, contrary to what we observed in our experiment. This is unsurprising, as there is nothing in the definitions given that may lead to ambiguity.

5.3 Existential readings of bathroom sentences (Elliott, 2020)

We may wonder whether an alternative definition of disjunction would lead to existential truth-conditions. Elliott (2020), the third work we'll look at, provides such a definition. To simplify[15], Elliott (2020)'s system can be seen as replacing the rule for disjunction in (30b) with (32b). This rule states that a positive update of *S or S′* is either (a) a positive update of *S* followed by a negative update of *S′* or (b) a negative update of *S* followed by a positive update of *S′* or (c) a positive update of *S* followed by a positive update of *S′*.

[15] Elliott does not consider just positive and negative updates but also undefined updates. Concomitantly, the rule for disjunction is a bit more complex than that presented in (32b). Another difference is that Elliott uses updates tagged with (trivalent) truth-values ; unlike Krahmer and Muskens (1995), the notions of positive and negative updates are not primitive, but defined as those updates from g to g', which are tagged with true and false respectively.

(32) a. $[\![S \text{ and } S']\!]^+ = \lambda g.\lambda g'.\ \exists h, [\![S]\!]^+(g)(h) \wedge [\![S']\!]^+(h)(g')$

b. $[\![S \text{ or } S']\!]^+ = \lambda g.\lambda g'.\ \begin{array}{r}(\exists h, [\![S]\!]^+(g)(h) \wedge [\![S']\!]^-(h)(g')) \\ \vee(\exists h, [\![S]\!]^-(g)(h) \wedge [\![S']\!]^+(h)(g')) \\ \vee(\exists h, [\![S]\!]^+(g)(h) \wedge [\![S']\!]^+(h)(g'))\end{array}$

The main motivation for Elliott (2020)'s proposal is that the denotation in (32b) is not arbitrary but can be shown to follow from the traditional Boolean disjunction via a systematic recipe. The existence of this recipe, Elliott argues, partially addresses Soames (1989)'s argument that dynamic theories are not predictive.

The reader can confer to Elliott (2020) for the general recipe but it is apparent in the definitions of (32): each clause in the definition of the positive update *S connective S′* corresponds to a case where *S connective S′* would be true in its classical semantics. Conjunction is only true when both its arguments is true and so it only has one such clause. Disjunction is true in three cases (first argument true second false, second argument true first false, both true) and so its positive update is the disjunction of three clauses corresponding to each of these cases.

As announced, the rule for disjunction in (32b) delivers existential truth conditions. To see this, consider a scenario where there are two circles, one blue, one green. This scenario makes the existential reading true and the universal reading false. In such a scenario, the negative update *there isn't a circle* would map the input assignment to two output assignments h_1 and h_2 (corresponding to each circle). The positive update of *it is green* would fail on h_2 (because the circle at g_i is blue not green) but succeed on h_1. So the update from g to h_1 would be possible ; in other words, the sentence would be true.

(33) $g = []$
$h_1 = [i \rightarrow$ green circle$]$
$h_2 = [i \rightarrow$ blue circle$]$
$g' = h_1 = [i \rightarrow$ green circle$]$

Problematically, the existential reading is the only one generated. As with Krahmer and Muskens (1995), there is no place in the theory where ambiguity might be generated. So this theory, as stated, is not compatible with our experimental results either.

While we highlighted the predictions for Elliott (2020), a number of other existing theories make the exact same predictions regarding cross-conjunction anaphora and bathroom sentences. In the interest of space, we summarize how they derive their predictions briefly.

Hofmann (2019, 2022) 8.name system assumes that a simple clause like "*there is a circle*" introduces a discourse referent at the highest scope possible ; this discourse referent is an individual concept defined in every world where there is circle and denoting a circle in those worlds. The sentence then states conditions on this discourse referent. In the case of conjunction, it imposes that the world of evaluation is one where there is a circle and where the circled picked by the individual concept discourse referent is red. In the case of disjunction, it imposes that either the world of evaluation is not one where there is a circle or the circled picked by the individual concept is red.

As in any dynamic system, the sentence is true if and only if there is *some* assignment of values to discourse referents which meets the conditions imposed on discourse referents. Put together, these assumptions imply that Hofmann predicts existential readings for both conjunctions and disjunctions ; indeed, the sentence will be true if one can find *some* individual concept that meets the requirements imposed by either sentence.

Mandelkern (2022) is a unique system aiming to derive discourse anaphora in purely static system. In this system, Sentences receive (in a world w and against an assignment g) truth-values and these truth values are derived using the entirely standard rules of first-order logic. Formulas receive, in addition to truth-values, so-called "bounds", which project and are filtered presuppositions but have a different pragmatic status. For instance, our sentences of interest have the following truth-conditions and bound satisfaction condition:

(34) There is a circle and it is blue.

a. True at (w, g) iff $[\exists x, \text{circle}(w)(x)] \wedge \text{blue}(g(i))$

b. Bounds satisfied iff $[\exists x, \text{circle}(w)(x)] \rightarrow \text{circle}(g(i))$

(35) Either there isn't a circle or it is blue.

a. True at (w, g) iff $[\neg\exists x, \text{circle}(w)(x)] \vee \text{blue}(g(i))$

b. Bounds satisfied iff $[\exists x, \text{circle}(w)(x)] \rightarrow \text{circle}(g(i))$

A sentence may be true or false and it may, independently, have its bounds satisfied (satt) or not (non-satt). Mandelkern (2022) isn't explicit on how his system derives the intuitions about the truth value of the sentence in a given scenario w *simpliciter*. Sentences only receive truth values when both a world and an assignment function are present. Some "*bridge assumption*" is needed to interpret the results of truth value judgment tasks such as ours. One plausible "*bridge assumption*"[16] would be: participants judge the sentence a truthful description of a picture if the sentence is true and has its bounds

satisfied in the world w represented by the picture for *some* assignment function g. This predicts existential judgments for both conjunction and disjunctions, as below.

(36) a. (34) judged true at w iff $\exists g, [\exists x, \text{circle}(w)(x)] \wedge \text{blue}(g(i)) \wedge [\exists x, \text{circle}(w)(x)] \rightarrow \text{circle}(g(i))$

b. (35) judged true at w iff $\exists g, [\neg\exists x, \text{circle}(w)(x)] \vee \text{blue}(g(i)) \wedge [\exists x, \text{circle}(w)(x)] \rightarrow \text{circle}(g(i))$

Alternative formulations of bridge assumption won't help. It is clear that every rule one might design will treat conjunction and disjunction configurations equally. Like previous theories reviewed, this system does not make a cut between the two types of configuration.

Finally, Gotham (2019) predicts explicit designs a system inspired by Krahmer and Muskens (1995) modified to derive uniqueness readings for bathroom sentences. If, this reading exists, which our experiments don't detect, it sits alongside the existential and universal readings. The presence of both of these readings is not predicted.

5.4 Connection to results on donkey anaphora

Despite all existing theories failing to derive our observation, our results are, in a certain sense, expected. In the small experimental literature devoted to donkey pronouns (Denić and Sudo, 2022; Foppolo, 2008; Geurts, 2002; Sun et al., 2020), it is found that, in experimental settings without particular context biases, the universal donkey sentences in (37b) is ambiguous between an existential and a universal reading. Existential sentences like (37a) are, on the other hand, non-ambiguous and only have existential readings.

(37) a. Some farmer who owns a donkey cherishes it.
$\exists x, [\ldots] \wedge [\ldots]$

b. Every farmer who owns a donkey cherishes it.
$\forall x, [\ldots] \rightarrow [\ldots]$

[16] This rule is probably what Mandelkern (2022) had in mind. Indeed, Mandelkern (2022) defines context as set of world-assignment pairs, states explicit update rules and defines a notion of null context. In that context, participants might be updating the null context with the sentence and judge it true if the picture corresponds to a world w which is the first component of a pair (w, g) in the updated context. Given Mandelkern (2022)'s definitions, that would be equivalent to the pragmatic rule given here.

Except for the quantifier, our sentences have similar logical structures to the ones considered in the donkey literature. Like universal donkey sentences, the bathroom sentences in (38b) can be expressed as a material conditional (→). Like existential donkey sentences, conjunctions sentences involve a conjunction of two basic clauses.

(38) a. There is a circle and it is blue.
$[\ldots] \wedge [\ldots]$

b. Either there isn't a circle or it is blue.
$\neg[\ldots] \vee [\ldots]$
$[\ldots] \rightarrow [\ldots]$

(39) a. Some farmer that has a donkey pats it.
$\exists x, [\ldots] \wedge [\ldots]$

b. No farmer that has a donkey pats it.
$\neg\exists x, [\ldots] \wedge [\ldots]$

c. Every farmer that has a donkey pats it.
$\forall x, [\ldots] \rightarrow [\ldots]$

We can phrase this more generally in terms of monotonicity following Kanazawa (1994): while existential readings are always available, universal readings are only available when they give rise to monotonic readings with respect to the antecedent's domain. For instance, under their universal reading, (40b) entails (40a). Likewise in the universal case, (41b) entails (41a). On the other hand, the universal readings of (42b) and (42a) are logically independent and likewise for the existential sentence in (43).

(40) a. Either there isn't a donkey or it is gray.

b. Either there isn't an animal or it is gray.

(41) a. Every farmer who has a donkey pats it.

b. Every farmer who has an animal pats it.

(42) a. There is a donkey and it is gray.

b. There is an animal and it is gray.

(43) a. Some farmer who has a donkey pats it.

b. Some farmer who has an animal pats it.

Equivalently, we could also express the generalization as follows: when the indefinite antecedent and the pronoun are in environments with the same monotonicity, only the existential reading is observed ; when they are in environments of opposite monotonicity, both readings are observed.

Regardless of the exact way of presenting the generalization, there seems to be parallels between what is observed in donkey (i.e. quantified) cases and non-quantified cases. This fact puts pressure on theories that locate the source of the existential/universal ambiguity in the quantifier. Indeed, an idea originating from Chierchia (1992), and also found in Champollion et al. (2017), is that quantifiers may be shifted from a traditional static meaning to two dynamic denotations, one that delivers the existential reading and the universal reading. As seen here, it seems that the rule that decides which type-shifter is available in the quantified case is the same as the one that decides which one is selected in the connective case. We do not claim this is a strong objection, but simply that our results demand more generality from the theories cited above.

6 Conclusion

In this paper, we presented evidence that sentences with cross-conjunction anaphora receive existential truth-conditions and bathroom sentences receive both existential and universal truth-conditions. This result, we argued, challenges established theories of these anaphora. While modifications to these theories may predict our results, they come at the cost of losing explanatoriness. Our results furthermore suggest a parallel between donkey anaphora and cross-connective anaphora. This suggests an uniform approach to pronoun readings in both configurations is needed.

References

Agresti, A. (2012). *Categorical Data Analysis*, volume 792. John Wiley & Sons.

Champollion, L., Bumford, D., and Henderson, R. (2017). Homogeneity in donkey anaphora. (2016):1–38.

Champollion, L., Bumford, D., and Henderson, R. (2019). Donkeys under discussion. *Semantics and Pragmatics*, 12:1–EA.

Chatain, K. (2018). Gaps in the interpretation of pronouns. In *Semantics and Linguistic Theory*, volume 28, pages 177–196.

Chierchia, G. (1992). Anaphora and dynamic binding. *Linguistics and philosophy*, 15:111–183.

Chierchia, G. (2009). *Dynamics of Meaning: Anaphora, Presupposition, and the Theory of Grammar*. University of Chicago Press.

Cooper, R. (1979). The interpretation of pronouns. *Syntax and Semantics*, 10:61–92.

Denic, M. (2020). Probabilistic informativeness in implicature computation: The case of embedded disjunctions.

Denić, M. and Sudo, Y. (2022). Donkey anaphora in non-monotonic environments. *Journal of Semantics*, 39(3):443–474.

Elliott, P. D. (2020). Towards a principled logic of anaphora.

Elliott, P. D. and Sudo, Y. (2023). Free choice with anaphora.

Evans, G. (1977). Pronouns, quantifiers, and relative clauses (I). *Canadian journal of philosophy*, 7(3):467–536.

Evans, G. (1980). Pronouns. *Linguistic Inquiry*, 11(2):337–362.

Foppolo, F. (2008). The puzzle of donkey anaphora resolution. In Schardl, A., Walkow, M., and Abdurrahman, M., editors, *North East Linguistics Society (NELS)*, volume 38, pages 297–310. GLSA.

Foster, E. D. and Deardorff, A. (2017). Open science framework (OSF). *Journal of the Medical Library Association: JMLA*, 105(2):203.

Geurts, B. (2002). Donkey Business. *Linguistics and Philosophy*, 25(2):129–129.

Gotham, M. (2019). Double Negation, Excluded Middle and Accessibility in Dynamic Semantics. (1991):1–10.

Groenendijk, J. and Stokhof, M. (1990). Dynamic montague grammar.

Groenendijk, J. and Stokhof, M. (1991). Dynamic predicate logic. *Linguistics and Philosophy*, 14(1):39–100.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis.

Hofmann, L. (2019). The anaphoric potential of indefinites under negation and disjunction. In *Proceedings of the 22nd Amsterdam Colloquium*, pages 181–190.

Hofmann, L. (2022). *Anaphora and Negation*. PhD thesis, University of California, Santa Cruz.

Kadmon, N. (1990). Uniqueness. *Linguistics and Philosophy*, 13(3):273–324.

Kamp, H., Van Genabith, J., and Reyle, U. (2011). Discourse representation theory. In *Handbook of Philosophical Logic*, pages 125–394. Springer.

Kanazawa, M. (1994). Weak vs. strong readings of donkey sentences and monotonicity inference in a dynamic setting. *Linguistics and Philosophy*, 17(2):109–158. NULL.

Krahmer, E. and Muskens, R. (1995). Negation and disjunction in discourse representation theory. *Journal of Semantics*, 12(4):357–376.

Lewis, D. (1975). Adverbs of quantification. *Formal semantics-the essential readings*, 178:188.

Mandelkern, M. (2020). Witnesses.

Mandelkern, M. (2022). Witnesses. *Linguistics and Philosophy*, 45(5):1091–1117.

Marty, P., Chemla, E., and Spector, B. (2015). Phantom readings: The case of modified numerals. *Language, Cognition and Neuroscience*, 30(4):462–477.

Muskens, R. (1996). Combining Montague Semantics and Discourse Representation. *Linguistics and philosophy*, 19(2):143–186. NULL.

Palan, S. and Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17:22–27.

Roberts, C. (1987). *Modal Subordination, Anaphora, and Distributivity*. PhD thesis.

Sauerland, U. (2003). A new semantics for number. In *Semantics and Linguistic Theory*, volume 13, pages 258–275.

Schubert, L. K. and Pelletier, F. J. (1989). Generically speaking, or, using discourse representation theory to interpret generics. *Properties, types and meaning*, 2:193–268.

Soames, S. (1989). Semantics and Semantic Competence. *Philosophical Perspectives*, 3:575–596.

Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In *Presupposition and Implicature in Compositional Semantics*, pages 243–281. Springer.

Sun, C., Rothschild, D., and Breheny, R. (2020). Exploring the existential/universal ambiguity in singular donkey sentences. In *Proceedings of Sinn Und Bedeutung*, volume 24, pages 289–305. PKP.

van der Does, J. (1993). The dynamics of sophisticated laziness. *Plurals and Anaphora. Dyana-2 Deliv*, pages 1–52.