Donkey Anaphora in Non-Monotonic Environments*

Milica Denić[1] and Yasutada Sudo[2]

[1] Universiteit van Amsterdam, Amsterdam, Netherlands
[2] University College London, London, UK

1 Introduction

Donkey anaphora is a type of pronominal anaphora involving a pronoun that is semantically bound by a non-c-commanding indefinite in a quantificational context as in (1).

(1) Every farmer who owns a donkey loves it.

We will call a pronoun involved in donkey anaphora a *donkey pronoun*.

Donkey pronouns are known to have two possible readings, the ∃-reading and the ∀-reading [Bra07, CBH18, Chi95, Fop08, Geu02, Kan94, Kri96, SP89, Yoo94, Yoo96].[1] For instance, without a context, (1) naturally receives a ∀-reading, (2-a), but as [Chi95] points out, in certain contexts, a weaker ∃-reading, (2-b), becomes available.

(2) a. Every donkey-owning farmer loves *all* of their donkeys.
b. Every donkey-owning farmer loves *at least one* of their donkeys.

The ∃-reading is observed with examples that make the ∀-reading implausible to be true due to world knowledge, as in (3) [Chi95, Kan94, SP89].

(3) a. Every man who had a quarter put it in the parking meter.
b. Every man who had a credit card paid his bill with it.

In addition to context, the quantifier of a donkey sentence is known to affect the perceived reading [Chi95, Kan94, Yoo94, Yoo96, CBH18]. Generally, donkey pronouns in the scope of universal quantifiers tend to receive ∀-readings, while those in the scope of *no* and existential quantifiers like *some* preferentially receive ∃-readings.

(4) No farmer who owns a donkey hates it.
≈ No donkey-owning farmer hates *any* of their donkeys.

(5) Some farmers who own a donkey love it.
≈ Some donkey-owning farmers love *at least one* of their donkeys.

Donkey anaphora involving monotonic quantifiers like the examples we've seen so far is relatively well discussed. In particular, [Kan94] proposes the following generalization (see [Yoo94, Yoo96, Geu02, Fop08] for experimental support).

*We wish to thank Emmanuel Chemla for many important discussions and suggestions. We also thank Keny Chatain, Gennaro Chierchia, Benjamin Spector and the audiences of LINGUAE seminar at ENS and XPRAG 2019 in Edinburgh for useful feedback and discussion. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 313610 and n. STG 716230 CoSaQ, and was supported by ANR-17-EURE-0017.

[1]For some speakers (1) is associated with a uniqueness presupposition that every farmer who owns a donkey owns only one, and with this presupposition the two readings collapse to one reading. Our experimental results indicate that this is not obligatory for our participants. See [Hei82, Roo87, Kad87, Hei90, Kad90, Kri96, Chi95, CBH18] for related discussion.

(6) a. Default readings for $\uparrow$MON$\uparrow$ quantifiers (e.g. *a, some*) and $\downarrow$MON$\downarrow$ quantifiers (e.g. *no*) are $\exists$-readings.
b. Default readings for $\uparrow$MON$\downarrow$ quantifiers (e.g. *all*) and $\downarrow$MON$\uparrow$ quantifiers (e.g. *not all*) are $\forall$-readings.

Compared to donkey anaphora involving monotonic quantifiers, the behavior of donkey pronouns in the scope of non-monotonic quantifiers is less well understood. For example, *most* and other proportional quantifiers are non-monotonic with respect to the NP argument, and [Kan94, 132] notes that donkey pronouns under their scope have no noticeable preferences between the two readings (see also [Hei82, §2.1.2] for a similar remark). On the other hand, [Kri96, 151] seems to think otherwise and remarks that proportional quantifiers do often give rise to $\forall$-readings.

[Kan94] also points out that 'existential non-monotonic quantifiers' like *exactly three*, which are non-monotonic with respect to both arguments, seem to preferentially receive $\exists$-readings.

(7) Exactly three farmers who own a donkey love it.
$\approx$ Three donkey-owning farmers love at least one of the donkeys they own and no other donkey-owning farmers love any of the donkeys they own.

More recently [CBH18] put forward an alternative theory to [Kan94] that (modulo contextual factors) predicts a 'conjunctive reading' for sentences like (7), which is essentially the conjunction of the existential and $\forall$-readings. Thus, under this view, (7) is purported to be true if and only if there are exactly three donkey-owning farmers who love their donkeys, and they all love all of their donkeys.

To our knowledge, no controlled empirical study so far has investigated the readings of donkey pronouns in non-monotonic environments, but as the above short review of the literature reveals, they are of particular theoretical interest, given that conflicting judgments have been reported and at the same time, different theories make different predictions.

The central aim of the present paper is to fill in this empirical gap by reporting on an experimental study that compares donkey anaphora involving two non-monotonic quantifiers, *exactly three* and *all but one*, using truth-value judgment tasks. The results of our experiments suggest that donkey anaphora involving *all but one* receives $\forall$-readings more prominently than donkey anaphora involving *exactly three*, for which we obtained no evidence that a $\forall$-reading can be accessed. In addition, we observe that the $\exists$-reading is easily accessible with both non-monotonic quantifiers, but do not find conclusive evidence that the aforementioned 'conjunctive' reading described by [CBH18] is accessed for either of the two quantifiers.

2 Experiment 1: *Exactly Three*

2.1 Design and Items

The meaning of a non-monotonic quantifier can be decomposed into an upward monotonic and a downward monotonic component. For instance, *exactly three* is semantically equivalent to the conjunction of the upward entailing quantifier *at least 3* and the downward entailing quantifier *at most 3*. Given this, there are in principle four logically possible readings for a donkey pronoun in the scope of *exactly three*: the pronoun could get (i) an $\exists$-reading in both components, (ii) a $\forall$-reading in both components, and (iii) a $\forall$-reading in the upward component combined with an $\exists$-reading in the downward component, and (iv) an $\exists$-reading in the upward component combined with the $\forall$-reading in the downward component. We will label these readings as

(i) $\frac{\exists}{\exists}$, (ii) $\frac{\forall}{\forall}$, (iii) $\frac{\forall}{\exists}$, (iv) $\frac{\exists}{\forall}$, respectively. The mnemonic is that what's above the line is the reading of the upward component of the meaning and what's below the line is the reading of the downward component of the meaning. For example, the four logically possible readings for (8) are paraphrased below.

(8) Exactly three squares that are above a heart are connected to it.

- $\frac{\exists}{\exists}$ At least three squares that are above a heart(s) are connected to **some** of those hearts and at most three squares that are above a heart(s) are connected to **some** of those hearts.
- $\frac{\forall}{\forall}$ At least three squares that are above a heart(s) are connected to **all** of those hearts and at most three squares that are above a heart(s) are connected to **all** of those hearts.
- $\frac{\forall}{\exists}$ At least three squares that are above a heart(s) are connected to **all** of those hearts and at most three squares that are above a heart(s) are connected to **some** of those hearts.
- $\frac{\exists}{\forall}$ At least three squares that are above a heart(s) are connected to **some** of those hearts and at most three squares that are above a heart(s) are connected to **all** of those hearts.

These four possible readings stand in the entailment relation depicted in Figure 1.

$$\begin{tikzcd}
 & \frac{\exists}{\exists} \arrow[dr, Rightarrow] & \\
\frac{\forall}{\exists} \arrow[ur, Rightarrow] \arrow[dr, Rightarrow] & & \frac{\exists}{\forall} \\
 & \frac{\forall}{\forall} \arrow[ur, Rightarrow] &
\end{tikzcd}$$

Figure 1: Entailments among the four logically possible readings

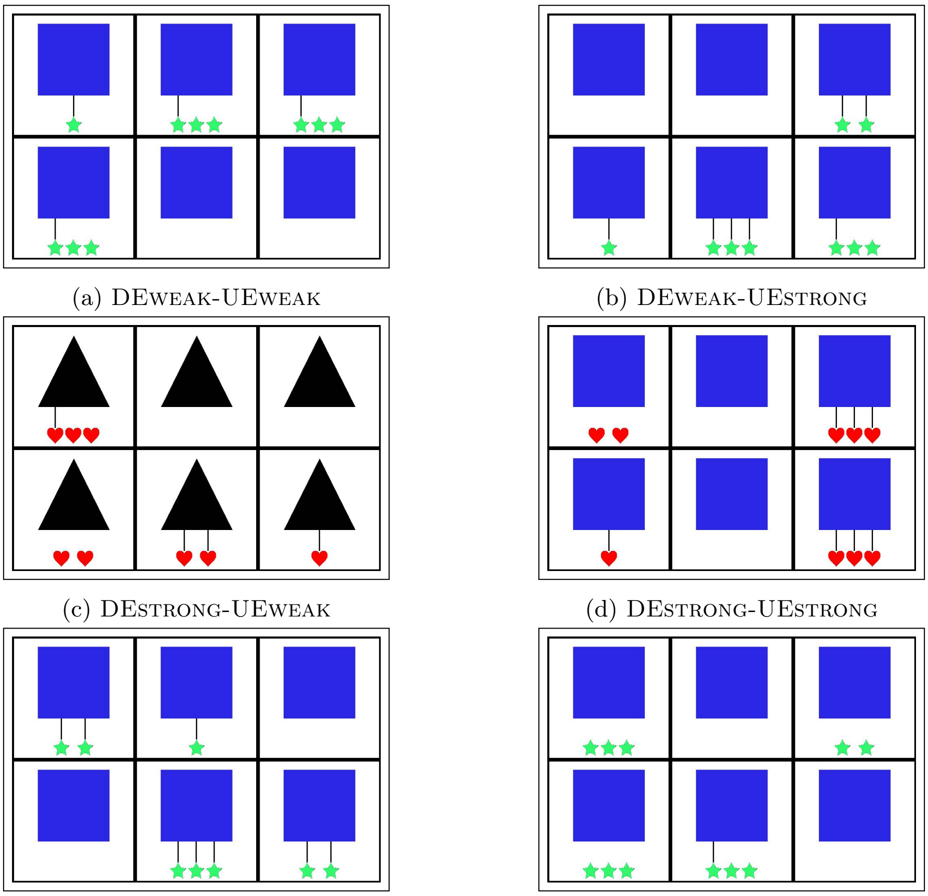
Test sentences in Experiment 1 were always of the following form:

(9) Exactly three of the ⟨squares, triangles⟩ that are above a ⟨star, heart⟩ are connected to it.

Given the entailment relation among the four readings, there are four kinds of situations where at least one of the readings is true, which constitute the target conditions of Experiment 1.

- DEweak-UEweak: Only the weakest reading $\frac{\exists}{\forall}$ is true, e.g. Figure 2a.
- DEweak-UEstrong: Only $\frac{\forall}{\forall}$ and $\frac{\exists}{\forall}$ are true, e.g. Figure 2b.
- DEstrong-UEweak: Only $\frac{\exists}{\exists}$ and $\frac{\exists}{\forall}$ are true, e.g. Figure 2c.
- DEstrong-UEstrong: All four readings are true, e.g. Figure 2d.

In addition there were two control conditions in the experiment, where none of the four readings was true. We refer to them as DEfalse-UEstrong and DEstrong-UEfalse: DEfalse-UEstrong is a type of context that makes the upward entailing part of the test sentence true under the ∀-reading, but falsifies the downward entailing part of the quantifier under both readings (e.g. Figure 2e). DEstrong-UEfalse makes the downward-entailing part of the test sentence true under the ∃-reading, but falsifies the upward entailing part of the quantifier under both readings (e.g. Figure 2f).



(a) DEweak-UEweak (b) DEweak-UEstrong

(c) DEstrong-UEweak (d) DEstrong-UEstrong

(e) DEfalse-UEstrong (f) DEstrong-UEfalse

Figure 2: Examples of experimental items.

This amounts to a total of six conditions (four target and two control conditions). Each of the six conditions had six items, meaning there were 36 test items.

Each image consisted of six vignettes as in Figures 2. Each of the vignettes contained a large shape of the same kind (either triangles or squares). In four out of six vignettes the square/triangle was above one, two, or three instances of smaller shapes of the same kind (either stars or a hearts). There were thus two vignettes in which the square/triangle was not above any hearts/stars, which ensured the felicity of the relative clause (cf. test sentence in (9)). In at least one of the vignettes the square or the triangle would appear above exactly one heart or star: this was to ensure the felicity of the singular morphology on the indefinite noun. For each item, a combination of shapes was chosen randomly (i.e. squares+stars, squares+hearts, triangles+stars, triangles+hearts), and the positions of the two vignettes with squares/triangles with no stars/hearts below them were chosen randomly as well. Likewise, the exact number of stars/hearts (one, two, or three) that appeared below the four squares/triangles in an item was chosen randomly for each of the four squares/triangles for each item, granting however that at least one of the squares/triangles would be above exactly one star/heart for felicity reasons mentioned above. We opted for having four squares/triangles that are above a star/heart in all of the test conditions because this permitted us to use the exact same visual stimuli for this experiment as for Experiment 2 that tested *all but one*.

2.2 Procedure and Participants

Participants were directed to a web-based truth-value judgment task, hosted on Alex Drummond's Ibex platform for psycholinguistic experiments. They were told that they would see sentences paired with images and that their task was to decide whether the sentence was true with respect to the image with which it was paired. The responses were given on a bounded continuous scale, whose ends were labeled as 'Completely false' and 'Completely true'.

The Participants first saw two practice trials, one involving a true sentence and one involving a false sentence, accompanied by suggested responses. The purpose of these examples was to familiarize the participants with the task. They then began the test phase of the experiment, the first two items of which were identical to the two practice trials. These were then followed by the 36 test items, presented in a randomized order for each participant.

65 participants (21 females) were recruited on Amazon Mechanical Turk. One participant was excluded for not being a native speaker of English. We furthermore excluded those participants whose average judgment in the four test conditions combined was lower than their average judgment in the two control conditions combined. The logic behind this exclusion criterion is the following. If they were able to access at least one of the four aforementioned logically possible readings, this should suffice for them to judge the test conditions on average better than the control conditions. If they did not do so, they might have not understood the experimental task, or they were possibly only able to access the uniqueness reading which was not verified in any of the six conditions and hence was not relevant for our purposes (cf. footnote 1). This led to the exclusion of two additional participants. The remaining 62 participants were thus kept for the analyses.

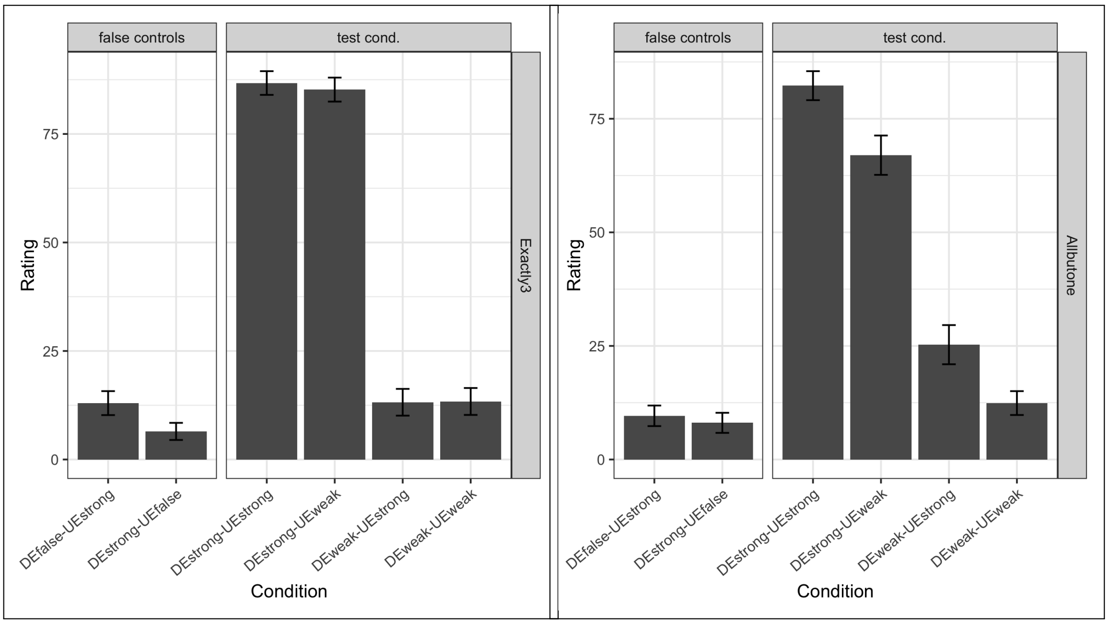
2.3 Results and Analysis

The results obtained are summarized in Figure 3a and Table 2. Recall that the target conditions render different logically possible readings true, as summarized in Table 1. Based on this, we will now discuss which readings the results give evidence for.

	$\frac{\forall}{\exists}$	$\frac{\exists}{\exists}$	$\frac{\forall}{\forall}$	$\frac{\exists}{\forall}$
DEstrong-UEstrong	T	T	T	T
DEstrong-UEweak	F	T	F	T
DEweak-UEstrong	F	F	T	T
DEweak-UEweak	F	F	F	T
DEfalse-UEstrong	F	F	F	F
DEstrong-UEfalse	F	F	F	F

Table 1: Target conditions and the readings that they make true.

No evidence for $\frac{\exists}{\forall}$ If the weakest reading, $\frac{\exists}{\forall}$, has been accessed, DEweak-UEweak, which validates $\frac{\exists}{\forall}$, should receive higher rating than the control conditions, which validate none of



(a) Experiment 1 (b) Experiment 2

Figure 3: Results of the two experiments per CONDITION. Error bars represent standard errors.

	Mean rating (SD)	
Condition	*exactly 3*	*all but one*
DEweak-UEweak	13.4 (3.1)	12.4 (2.6)
DEweak-UEstrong	13.2 (3.1)	25.3 (4.3)
DEstrong-UEweak	85.2 (2.8)	67 (4.3)
DEstrong-UEstrong	86.7 (2.7)	82.3 (3.2)
DEfalse-UEstrong	13 (2.8)	9.6 (2.3)
DEstrong-UEfalse	6.5 (2)	8.1 (2.2)

Table 2: Experiments 1 and 2: Mean participants' rating and standard deviation per condition.

the readings. The data was subsetted to the items in DEfalse-UEstrong and DEweak-UEweak conditions[2]. A linear mixed model was fitted on this data set with CONDITION as a fixed effect and random by-participant intercepts and slopes. A comparison of this model with a reduced model without CONDITION as a fixed effect revealed no significant effect of CONDITION ($\chi(1) = 0.06, p = .8$). There is thus no evidence for the existence of $\frac{\exists}{\forall}$ reading with *exactly 3*.

No evidence for $\frac{\forall}{\forall}$ If $\frac{\forall}{\forall}$ has been accessed, DEweak-UEstrong, which validates both $\frac{\forall}{\forall}$ and $\frac{\exists}{\forall}$, should receive higher rating than DEweak-UEweak, which only validates $\frac{\exists}{\forall}$. The data was subsetted to the items in DEweak-UEstrong and DEweak-UEweak conditions.

[2]DEfalse-UEstrong was chosen rather than DEstrong-UEfalse because the mean rating of DEfalse-UEstrong was higher than that of DEstrong-UEfalse, and thus provides a stricter requirement for the detection of $\frac{\exists}{\forall}$.

A linear mixed model was fitted on this data set with condition as a fixed effect and random by-participant intercepts and slopes. A comparison of this model with a reduced model without condition as a fixed effect revealed no significant effect of condition ($\chi(1) = 0.02, p = .9$). There is thus no evidence for the existence of $\substack{\forall\\\forall}$ reading with *exactly 3*.

Evidence for $\substack{\exists\\\exists}$ If $\substack{\exists\\\exists}$ has been accessed, DEstrong-UEweak, which validates both $\substack{\exists\\\exists}$ and $\substack{\exists\\\forall}$, should receive higher rating than DEweak-UEweak, which only validates $\substack{\exists\\\forall}$. The data was subsetted to items in DEstrong-UEweak and DEweak-UEweak conditions. A linear mixed model was fitted on this data set with condition as a fixed effect and random by-participant intercepts and slopes. A comparison of this model with a reduced model without condition as a fixed effect revealed a significant effect of condition ($\chi(1) = 100, p < .001$). Our results thus provide evidence for the existence of $\substack{\exists\\\exists}$ reading with *exactly 3*.

No evidence for $\substack{\forall\\\exists}$ In order to see whether the $\substack{\forall\\\exists}$ reading is available or not, we cannot simply compare DEstrong-UEstrong condition, which is the only condition validating $\substack{\forall\\\exists}$, to some other condition. The reason is the following. Suppose the reading $\substack{\forall\\\exists}$ is never accessed, while the other three readings (i.e. $\substack{\forall\\\forall}$, $\substack{\exists\\\exists}$, and $\substack{\exists\\\forall}$) are accessed at least to some extent. This would mean that DEstrong-UEstrong validates all of the three available readings, while all the other conditions validate at most a proper subset thereof. This on its own might suffice to make participants rate items in DEstrong-UEstrong condition higher than in any of the remaining conditions. Therefore, a significant difference between DEstrong-UEstrong and any of the other conditions in itself would not constitute strong evidence for the existence of $\substack{\forall\\\exists}$.

To circumvent this issue, we selected participants with the following property: their mean rating in at least one of DEstrong-UEweak and DEweak-UEstrong is equal or lower than in DEweak-UEweak. The idea is that these participants accessed at most one of $\substack{\exists\\\exists}$ and $\substack{\forall\\\forall}$. In other words, they (at most) accessed either (i) $\substack{\forall\\\forall}$, $\substack{\forall\\\exists}$, and $\substack{\exists\\\forall}$ or (ii) $\substack{\exists\\\exists}$, $\substack{\forall\\\exists}$, and $\substack{\exists\\\forall}$. This further means that, for the participants who did not access $\substack{\forall\\\forall}$, the only reading which is true in DEstrong-UEstrong but not in DEstrong-UEweak is $\substack{\forall\\\exists}$. Likewise, for participants who did not access $\substack{\exists\\\exists}$, the only reading which is true in DEstrong-UEstrong but not in DEweak-UEstrong is $\substack{\forall\\\exists}$. Thus, if these participants would rate DEstrong-UEstrong even better than the one they rated better between DEstrong-UEweak and DEweak-UEstrong, this could be taken as evidence that these participants accessed $\substack{\forall\\\exists}$. 42 participants fell into this category, and the following analysis was conducted on their responses.

The data was subsetted to items in DEstrong-UEstrong and the better rated condition between DEstrong-UEweak and DEweak-UEstrong (as determined for each participant separately). A linear mixed model was fitted on this data set with condition (DEstrong-UEstrong vs. Other) as a fixed effect and random by-participant intercepts and slopes. A comparison of this model with a reduced model without condition as a fixed effect revealed no significant effect of condition ($\chi(1) = 0.01, p = .9$). For reference, the mean rating of the better rated conditions between DEstrong-UEweak and DEweak-UEstrong (as determined for each participant separately) was 88.6 ($SD = 2.8$), while their mean rating in DEstrong-UEstrong was 88.7 ($SD = 3.4$). There is thus no evidence for the existence of $\substack{\forall\\\exists}$ with *exactly 3*.

3 Experiment 2: *All but One*

3.1 Design and Items

Experiment 2 had the exact same design as Experiment 1 except that the test sentences used *all but one*, in place of *exactly three* as in (10).

(10) All but one of the ⟨squares, triangles⟩ that are above a ⟨star, heart⟩ are connected to it.

As mentioned above, we constructed the pictures for Experiment 1 in such a way that they can be used in Experiment 2 as well. Thus, any differences between the results of the two experiments have to be due to the linguistic, rather than visual, stimuli.

3.2 Procedure and Participants

The procedure was identical to Experiment 1. A new set of 65 participants (25 females) were recruited on Amazon Mechanical Turk, none of whom participated in Experiment 1. One participant was excluded for failing to complete the experiment, two participants were excluded for not being native speakers of English, and six participants were excluded for their average judgment in target conditions not being higher than their average judgment in control conditions (which is the same exclusion criterion as in Experiment 1). 56 participants were thus kept for the analysis.

3.3 Results and Analysis

The results obtained are summarized in Figure 3b and Table 2. The logic of the data analysis is identical to Experiment 1, and we've conducted statistical analyses parallel to Experiment 1 as follows.

No evidence for $\genfrac{}{}{0pt}{}{\exists}{\forall}$ Statistical analyses on data from DEweak-UEweak and DEfalse-UEstrong revealed no significant effect of condition ($\chi(1) = 2.35, p = .12$). There is thus no evidence for the existence of $\genfrac{}{}{0pt}{}{\exists}{\forall}$ with *all but one.*

Evidence for $\genfrac{}{}{0pt}{}{\forall}{\forall}$ Statistical analyses on data from DEweak-UEweak and DEweak-UEstrong showed that unlike in Experiment 1, DEweak-UEstrong was judged significantly better than DEweak-UEweak ($\chi(1) = 10.5, p < .01$). This result provides evidence for the existence of $\genfrac{}{}{0pt}{}{\forall}{\forall}$ with *all but one.*

Evidence for $\genfrac{}{}{0pt}{}{\exists}{\exists}$ Statistical analyses on data from DEweak-UEweak and DEstrong-UEweak indicate that as in Experiment 1, DEstrong-UEweak is judged significantly better than DEweak-UEweak ($\chi(1) = 64.8, p < .001$). This result provides evidence for the existence of $\genfrac{}{}{0pt}{}{\exists}{\exists}$ with *all but one.*

Weak evidence for $\genfrac{}{}{0pt}{}{\forall}{\exists}$ As in Experiment 1, in order to determine whether participants have accessed $\genfrac{}{}{0pt}{}{\forall}{\exists}$, we selected participants whose mean rating in at least one of DEstrong-UEweak and DEweak-UEstrong is equal or lower than in DEweak-UEweak. 33 participants fell into this category in Experiment 2. Analyses parallel to Experiment 1 were conducted on their responses in DEstrong-UEstrong and the better rated condition between

DEstrong-UEweak and DEweak-UEstrong (as determined for each participant separately). They revealed a borderline effect of condition (DEstrong-UEstrong vs. other) ($\chi(1) = 3.62, p = .057$). For reference, the mean rating of the better rated conditions between DEstrong-UEweak and DEweak-UEstrong (as determined for each participant separately) was 73.2 ($SD = 5.6$), while the mean rating in DEstrong-UEstrong was 82.3 ($SD = 4.6$). This suggests that $\frac{\forall}{\exists}$ might be available with *all but one*, but further research is needed to establish this conclusively.

4 Concluding Remarks

To summarize the main empirical findings, the results of the two truth-value judgment experiments showed that the $\frac{\exists}{\exists}$ reading of donkey anaphora is available with both *exactly 3* and *all but one*, and we obtained no evidence that $\frac{\exists}{\forall}$ reading is available with either. Furthermore, we find some suggestive evidence that the $\frac{\forall}{\exists}$ reading might be available with *all but one*, but further data should be assessed before a firm conclusion could be drawn. Most interestingly, we find differences between *all but one* and *exactly 3* with respect to the availability of the $\frac{\forall}{\forall}$ reading.

As the logical monotonicity profiles of the two quantifiers are the same, something else must be behind the observed difference. A direction we explored in a follow-up experiment is that subjective, rather than logical, monotonicity might be what matters for the interpretation of donkey anaphora, as [CHR11] claim for NPI licensing. Focusing on quantifier *all*, we tested whether participants' perceived monotonicity of this quantifier explains the extent to which different readings of donkey anaphora are available with it. Due to the limited space available here, we have to omit details of this follow-up experiment, but we found no evidence that subjective monotonicity plays a role in donkey anaphora interpretation with quantifier *all*. It is thus unlikely that subjective perception of monotonicity properties of *all but one* and *exactly 3* are to explain the differences between the two quantifiers in terms of donkey anaphora interpretation.

In light of this, we see two other promising directions to be explored in future research. The first is the possibility that the symmetry profile of the quantifier affects the default reading, as suggested by [Kan94].[3] That is, *exactly 3* is symmetric, but *all but one* is not. The second possibility connects to the well-known fact that context and question under discussion can influence the preferred reading of donkey pronouns (cf. [Chi95, CBH18]). The idea is that *all but one* and *exactly 3*, despite having the same monotonicity properties, might be typically used to answer different questions under discussion, which results in the differences in their default readings.

References

[Bra07] Adrian Brasoveanu. *Structured Nominal and Modal Reference.* Ph.D. dissertation, Rutgers University, 2007.

[CBH18] Lucas Champollion, Dylan Bumford, and Robert Henderson. Donkeys under discussion. To appear in Semantics and Pragmatics, 2018.

[Chi95] Gennaro Chierchia. *Dynamics of Meaning: Anaphora, Presupposition, and the Theory of Grammar.* University of Chicago Press, Chicago, 1995.

[3] [Kan94] also mentions Left Continuity as a potentially relevant property here, but since both *exactly 3* and *all but one* are left continuous, it wouldn't explain the difference we observed.

[CHR11] Emmanuel Chemla, Vincent Homer, and Daniel Rothschild. Modularity and intuitions in formal semantics: the case of polarity items. *Linguistics and Philosophy*, 34(6):537–570, 2011.

[Fop08] Francesca Foppolo. The puzzle of donkey anaphora resolution. In *NELS*, 2008.

[Geu02] Bart Geurts. Donkey business. *Linguistics and Philosophy*, 25(2):129–156, 2002.

[Hei82] Irene Heim. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. dissertation, University of Massachusetts, Amherst, 1982.

[Hei90] Irene Heim. E-type pronouns and donkey anaphora. *Linguistics and Philosophy*, 13(2):137–177, 1990.

[Kad87] Nirit Kadmon. *On unique and non-unique reference and asymmetric quantification*. Ph.D. dissertation, University of Massachusetts, Amherst, 1987.

[Kad90] Nirit Kadmon. Uniqueness. *Linguistics and Philosophy*, 13(3):273–324, 1990.

[Kan94] Makoto Kanazawa. Weak vs. strong readings of donkey sentences and monotonicity inferences in a dynamic setting. *Linguistics and Philosophy*, 17(2):109–158, 1994.

[Kri96] Manfred Krifka. Pragmatic strengthening in plural predications and donkey sentences. In *Proceedings of SALT 6*, pages 136–153, 1996.

[Roo87] Mats Rooth. Noun Phrase Interpretation in Montague Grammar, File Change Semantics, and Situation Semantics. In Peter Gärdenfors, editor, *Generalized Quantifiers: Linguistic and Logical Approaches*, pages 237–268. Reidel, Dordrecht, 1987.

[SP89] Lenhart K. Schubert and Francis Jeffery Pelletier. Generically speaking, or, using discourse representation theory to interpret generics. In Gennaro Chierchia, Barbara H. Partee, and Raymond Turner, editors, *Properties, Types and Meaning II: Semantic Issues*, pages 193–268. Kluwer Academic Publishers, 1989.

[Yoo94] Youngeun Yoon. *Weak and Strong Interpretations of Quantifiers and Definite NPs in English and Korean*. PhD thesis, University of Texas at Austin, 1994.

[Yoo96] Youngeun Yoon. Total and partial predicates and the weak and strong interpretations. *Natural Language Semantics*, 4(3):217–236, 1996.