Donkey anaphora in non-monotonic environments: An experimental investigation

Milica Denić (ENS) and Yasutada Sudo (UCL)

Donkey anaphora in quantified sentences such as (1) is ambiguous between existential (1a) and universal reading (1b). Which reading is preferred depends on the quantifier: universally quantified sentences like (1) favor the universal reading, while existentially quantified sentences favor the existential reading (see Geurts 2002, Foppolo 2008 for experimental data). The monotonicity profile of the quantifier is considered to be a major factor in deciding which reading is preferred (Kanazawa 1994, a.o.).

(1) Every farmer who owns a donkey loves it.
   a. Every farmer who owns a donkey loves some of his donkeys. **existential reading**
   b. Every farmer who owns a donkey loves all of his donkeys. **universal reading**

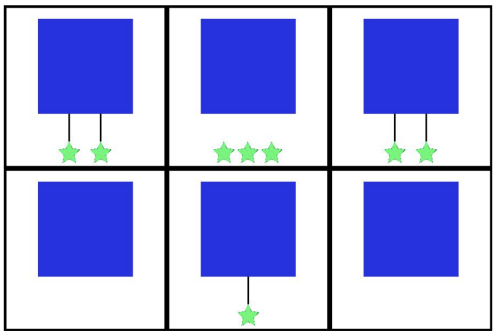
Non-monotonic (NM) quantifiers can be seen as simultaneously having an upward-entailing (UE) and downward-entailing (DE) component. For instance, a NM quantifier *exactly 3* is semantically equivalent to the conjunction of the UE quantifier *at least 3* with the DE quantifier *at most 3*. There are thus in principle four logically possible readings for donkey sentences with NM quantifiers. The pronoun could get (i) an existential reading in both components, (ii) a universal reading in both components, (iii) an existential reading in the UE component + universal reading in the DE component, and (iv) a universal reading in the UE component + existential reading in the DE component. Since the existential reading is stronger in the DE component but weaker in the UE component, we will label these logically possible readings as (i) DEstrong-UEweak, (ii) DEweak-UEstrong, (iii) DEweak-UEweak, (iv) DEstrong-UEstrong, respectively. For instance, the DEstrong-UEweak reading of (2a) is (2b).

(2) a. Exactly 3 squares that are above a heart are connected to it.
   b. At least 3 squares that are above a heart are connected to **some** of those hearts and at most three squares that are above a heart are connected to **some** of those hearts.

Judgments reported in the literature for donkey sentences with NM quantifiers are equivocal. For instance, Kanazawa (1994) reports what we call the DEstrong-UEweak reading for donkey anaphora under *exactly* n, and proposes two hypotheses that would explain this, while Champollion et al. (2018) contend that the main interpretation is the DEstrong-UEstrong reading for all NM quantifiers, and develop an account that derives it. In this study, we experimentally investigate which of the four logically possible readings of donkey anaphora are actually attested in the scope of two NM quantifiers: *exactly 3* and *every but one*. These two quantifiers have the same monotonicity profiles, and are truth-conditionally equivalent when the cardinality of the restrictor is 4. To summarize the findings: for both *exactly 3* and *every but one*, we find evidence for the DEstrong-UEweak reading but no evidence for the DEweak-UEweak or DEstrong-UEstrong reading; furthermore, there is evidence for the availability of the DEweak-UEstrong reading for *every but one* but not for *exactly 3*.

Experiment: The experiment employed a truth-value judgment task with binary responses (True or False). 72 participants were assigned to one of the two versions of the experiment, which differed only in the determiner of the test sentences, which looked like (3a) or (3b).

(3) a. Every ⟨square, triangle⟩ but one that is above a ⟨star, heart⟩ is connected to it.
   b. Exactly 3 ⟨squares, triangles⟩ that are above a ⟨star, heart⟩ are connected to it.



Every square but one that is above a star is connected to it.

Figure 1

Each participant saw 14 test items belonging to six conditions: (1) In DEweak-UEweak only the DEweak-UEweak reading was true; (2) In DEweak-UEstrong only the DEweak-UEweak and DEweak-UEstrong readings were true (the latter entails the former); (3) In DEstrong-UEweak only the DEweak-UEweak and DEstrong-UEweak readings were true (the latter entails the former); (4) In DEstrong-UEstrong all four readings were true (the DEstrong-UEstrong reading entails all the others); (5) DEfalse-UEstrong and (6) DEstrong-UEfalse were control conditions and didn't verify any logically possible readings. Figure 1 is an example item in DEstrong-UEstrong for the 'every but one' version of the experiment. There were four experimental items in DEstrong-UEstrong, and two experimental items in each of the five remaining con-

ditions. The distribution was chosen to ensure an equal number of items that were true and of items that are false on all logically possible readings.

Results: Logit mixed effects models were fit with CONDITION as a fixed effect and random by-participant intercepts, separately for the following pairs of conditions: (i) DEFALSE-UESTRONG and DEWEAK-UEWEAK (ii) DESTRONG-UEWEAK and DEWEAK-UEWEAK, (iii) DEWEAK-UESTRONG and DEWEAK-UEWEAK. (i) revealed no significant effect of CONDITION for either *every but one* ($\chi(1) = 1.03, p = .31$) or for *exactly 3* ($\chi(1) = 2.6, p = .11$): **there is thus no evidence for the existence of the DEweak-UEweak reading**. (ii) revealed a significant effect of CONDITION both for *every but one* ($\chi(1) = 27.06, p < .001$) and for *exactly 3* ($\chi(1) = 61.85, p < .001$): **there is thus evidence for the existence of the DEstrong-UEweak reading**. (iii) revealed a significant effect of condition for the quantifier *every but one* ($\chi(1) = 7.51, p < .01$): **there is thus evidence for the DEweak-UEstrong reading for the quantifier *every but one***. As the mean is lower in DEWEAK-UESTRONG than in DEWEAK-UEWEAK condition for *exactly 3* (cf. Figure 2), **there is no evidence for the DEweak-UEstrong reading with *exactly 3***. Answering the question of whether there is evidence for the DEstrong-UEstrong reading is more complicated. Since the DEstrong-UEstrong reading entails all the other readings, participants might have accessed only the DEstrong-UEweak and DEweak-UEstrong readings, and said true more often in DESTRONG-UESTRONG because all the readings accessed are true there, unlike in the other conditions where only a subset of them is true. To this end, we analyzed the data from the 50 participants who consistently said false in at least one of DESTRONG-UEWEAK and DEWEAK-UESTRONG. The idea is that these participants accessed at most one reading which is not the DEstrong-UEstrong reading. Logit mixed models on the results from DESTRONG-UESTRONG and whichever of DESTRONG-UEWEAK and DEWEAK-UESTRONG in which they said 'True' (if none, their responses were coded as 'False') revealed no significant effect of CONDITION for either *every but one* ($\chi(1) = 0.27, p = .6$) or for *exactly 3* ($\chi(1) = 1.08, p = .3$): **there is thus no evidence for the existence of the DEstrong-UEstrong reading**.

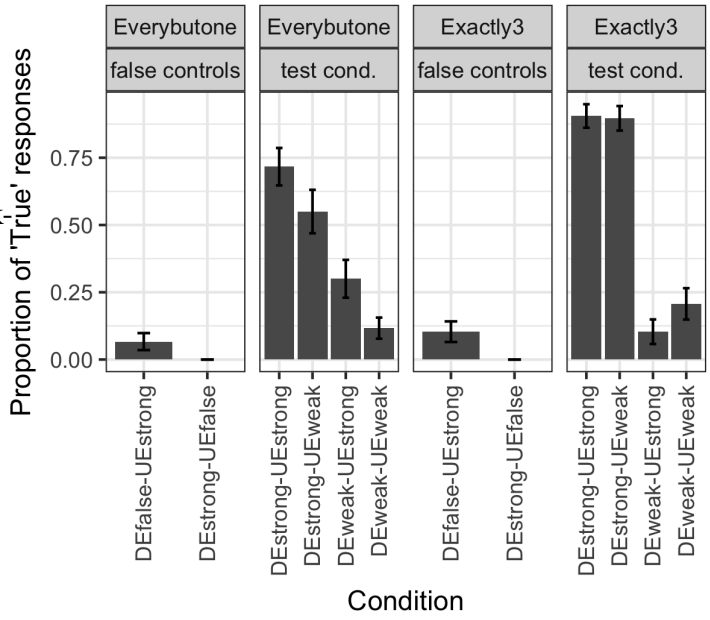


Figure 2

Discussion: We will discuss the following two theoretical consequences: (1) Implications of the lack of evidence for the DEstrong-UEstrong reading for Champollion et al. (2018). (2) The difference between *every but one* and *exactly 3* with respect to the DEweak-UEstrong reading suggests that the monotonicity profile is not the only major factor for determining the preferred reading (contrary to Kanazawa 1994). Regarding (2), we would like to point out an interesting theoretical possibility that would make these results compatible with Kanazawa's proposal. The idea is that subjective, rather than logical, monotonicity might be what matters for the interpretation of donkey anaphora, as Chemla et al. (2011) claim for NPI licensing. In other words, it is possible that people do not perceive *every but one* and *exactly 3* as having the same monotonicity profiles and that this in turn determines the extent to which different readings are accessed. This hypothesis is empirically testable, and we intend to address this question in future experimental work. There is an alternative to explore regarding (2), connecting to the well-known fact that context and question under discussion can influence the reading preferences (universal vs. existential) of donkey pronouns (cf. Chierchia 1992, and more recently Champollion et al. 2018). The idea is that *every but one* and *exactly 3*, despite having the same monotonicity profiles, are typically used to answer very different questions under discussion, and that this determines the extent to which the DEweak-UEstrong and the DEstrong-UEweak readings are accessible.

References: Champollion et al. (2018) Donkeys under discussion. *Sem & Prag*. Chierchia (1992) Anaphora and dynamic binding. *L&P*, 15. Foppolo (2008) The puzzle of donkey anaphora resolution. *NELS 38*. Geurts (2002) Donkey business. *L&P*, 25. Kanazawa (1994) Weak vs. strong readings of donkey sentences and monotonicity inferences in a dynamic setting. *L&P*, 17.