An experimental investigation of implicature and homogeneity approaches to free choice*

Lyn Tieu
University of Toronto

Cory Bill
Leibniz-ZAS

Jacopo Romoli
Heinrich-Heine Universität Düsseldorf

Abstract A sentence containing disjunction in the scope of a possibility modal, such as *Angie is allowed to buy the boat or the car*, gives rise to the FREE CHOICE inference that Angie can freely choose between the two. This inference poses a well-known puzzle, in that it is not predicted by a standard treatment of modals and disjunction (e.g., Kamp 1974). To complicate things further, FREE CHOICE tends to disappear under negation: *Angie is not allowed to buy the boat or the car* doesn't merely convey the negation of free choice, but rather the stronger DOUBLE PROHIBITION reading that Angie cannot buy either one. There are two main approaches to the FREE CHOICE-DOUBLE PROHIBITION pattern in the literature. While they both capture the relevant data points, they make a testable, divergent prediction regarding the status of positive and negative sentences in a context in which Angie can only buy one of the two objects, e.g., the boat. In particular, the implicature-based approach (e.g., Fox 2007, Klinedinst 2007, Bar-Lev & Fox 2017, 2020) predicts that the positive sentence is true in such a context, but associated with a false implicature, while it predicts the negative sentence to be straightforwardly false. The homogeneity-based approach in Goldstein (2019) predicts both the positive and negative sentences to be equally undefined (see also Aloni 2022 and Willer 2017 for similar predictions). Investigating the contrast between these sentences in such a context therefore provides a clear way to address the debate between implicature and non-implicature accounts of FREE CHOICE. We present a set of four experiments aiming to do just this, by comparing FREE CHOICE inferences to regular implicatures, using a ternary judgment task. The results overall present a challenge for the implicature approach. We discuss how the implicature approach could be amended to account for our results, based on a recent proposal by Enguehard & Chemla (2021) on the distribution of implicatures.

Keywords: free choice, implicature, homogeneity, ternary judgment task, polarity

* Parts of this manuscript were published in the Proceedings of the 29th Semantics and Linguistic Theory conference, specifically the introduction and an earlier version of Experiment 1.

1 Introduction

A sentence containing disjunction in the scope of a possibility modal, such as (1a), gives rise to the so-called FREE CHOICE inference in (1b), conveying that Angie is allowed to buy the boat and that she is allowed to buy the car. That is, she can *freely choose* between the two.

(1) a. Angie is allowed to buy the boat or the car.
b. ⇝ *Angie can choose between the two* FREE CHOICE

This inference poses a well-known puzzle, in that it is not predicted by standard treatments of modals and disjunction.[1] To complicate things further, FREE CHOICE tends to disappear under negation: the corresponding negative sentence in (2a) doesn't merely convey the negation of free choice; rather, it conveys the stronger DOUBLE PROHIBITION reading, conveying that Angie cannot buy either one.

(2) a. Angie is not allowed to buy the boat or the car.
b. ⇝ *Angie is not allowed to buy either one* DOUBLE PROHIBITION

There are two main approaches to the FREE CHOICE-DOUBLE PROHIBITION pattern in the literature. One is based on deriving FREE CHOICE as an implicature, while the other either encodes it in the meaning of disjunction or the modal, or treats it as a pragmatic inference of a different kind. Both approaches can account for the basic cases above, as well as a variety of more complex data. They differ, however, with respect to a simple prediction regarding the status of sentences like (1a) and (2a) in a context where Angie can only buy one of the two objects, e.g., the boat. In particular, the implicature approach predicts the positive case to be a true sentence with a false implicature, while it predicts the negative case to be straightforwardly false in such a context. The non-implicature approach, on the other hand, predicts both cases to be equally undefined. Comparing these sentences in such a context, therefore, provides a clear way to address the debate between implicature and non-implicature approaches to FREE CHOICE.

We present a set of four experiments aimed at testing these predictions, using a ternary judgment task (see Katsos & Bishop 2011, Abrusan & Szendroi 2013, Križ & Chemla 2015, Tieu et al. 2017a, Renans et al. 2018, Tieu et al. 2019, among others) to target undefinedness.

Experiment 1 focused on the comparison between free choice disjunction and plain disjunction as in (3a), the EXCLUSIVITY inference of which is less controversially analyzed as an implicature (3b).

1 See Kamp (1974) and much subsequent work, as well as Meyer (2018) for a detailed overview of the problem.

(3) a. Angie bought the boat or the car.
b. ⇝ *Angie didn't buy both the boat and the car* IMPLICATURE

Experiments 2a and 2b compared free choice disjunction to the *not required to* implicature of weak modals, as in (4).

(4) a. Angie is allowed to buy the boat.
b. ⇝ *Angie is not required to buy the boat* IMPLICATURE

In Experiment 3, we addressed potential confounds associated with Experiments 1, 2a, and 2b, and compared free choice 'any' to the quantifier 'some'.

(5) a. Angie is allowed to buy any of the food items.
b. ⇝ *Angie can choose between the food items* FREE CHOICE

(6) a. Angie bought some of the food items.
b. ⇝ *Angie didn't buy all of the food items* IMPLICATURE

To anticipate the results, all four experiments revealed a significant interaction between inference type and polarity, suggesting free choice and scalar implicatures do not pattern in the same way with respect to polarity. Along the way, however, we also find that the interpretation of the ternary judgment task has its own challenges that need to be controlled for. We discuss the methodological implications of our study, and, on the theoretical side, how the implicature approach could be amended to account for our results, based on a recent proposal by Enguehard & Chemla (2021) on the distribution of implicatures.

The rest of the paper is organized as follows. In Section 2, we provide further background on the FREE CHOICE-DOUBLE PROHIBITION pattern, outlining the two approaches and their predictions. In Section 3, we briefly discuss some relevant previous studies. We present our experiments in Sections 4, 5, and 8, and in Section 9 we discuss the implications of the findings for theories of free choice, including how the implicature approach might be further developed to account for the results. Section 10 concludes the paper.

2 Background

2.1 Free choice

As mentioned, a sentence like (1a) and its negation in (2a) pose a challenge for standard treatments of modals and disjunction. To illustrate the pattern more schematically, a configuration like (7a) gives rise to the conjunctive FREE CHOICE inference in (7b). Its negation in (8a) gives rise to a conjunctive DOUBLE PROHIBITION inference, which is stronger than the NEGATED FREE CHOICE inference (which would

correspond to: $\neg \Diamond B \vee \neg \Diamond C$).

(7) a. $\Diamond(B \vee C)$
b. $\Diamond B \wedge \Diamond C$ FREE CHOICE

(8) a. $\neg \Diamond (B \vee C)$
b. $\neg \Diamond B \wedge \neg \Diamond C$ DOUBLE PROHIBITION

In addition, the negated free choice reading re-emerges in cases like (9), in which a DOUBLE PROHIBITION reading of the first sentence would lead to a contradiction with the continuation. The coherence of (9) indicates that the first sentence is read with a negated free choice reading.

(9) Angie isn't allowed to buy the boat or the car. She's only allowed to buy the boat!

Relatedly, the FREE CHOICE inference can also be suspended in cases like (10), where a FREE CHOICE reading of the first sentence would lead to a contradiction with the second one. The coherence of (10) again tells us that the first sentence can be read as not entailing FREE CHOICE.

(10) Angie is allowed to buy the boat or the car. I don't remember which one.

Any approach to the FREE CHOICE-DOUBLE PROHIBITION pattern not only has to predict FREE CHOICE in the positive case and DOUBLE PROHIBITION in the negative one, but it also needs to account for the absence of such DOUBLE PROHIBITION and FREE CHOICE readings in cases like (9) and (10), respectively.[2] We turn now to the two main approaches in the literature.

2.2 Two approaches

2.2.1 The implicature approach

The most prominent approach in the literature is based on a theory of implicatures and comes in different versions (e.g., Kratzer & Shimoyama 2002, Fox 2007, Klinedinst 2007, Chemla 2010, Franke 2011, Santorio & Romoli 2017, Bar-Lev 2018, Bar-Lev & Fox 2020, Del Pinal et al. 2022). Without going into the details of the implementation, this account is based on three main ingredients: (i) a standard meaning for disjunction and possibility modals, (ii) an implicature-generating algorithm, which we can refer to as 'EXH', and (iii) a(n independently required) principle regulating the distribution of EXH, which bans or strongly disfavors EXH under

2 The puzzle, in fact, extends more generally beyond the positive/negative dichotomy, to upward- versus downward-entailing environments; see Fox (2007), among others, for discussion.

negation, as in (11) (see, among others, Chierchia et al. 2012 and Fox & Spector 2018).

(11) Do not insert EXH in a sentence *S* if the resulting meaning is weaker than that of *S*, unless forced to.

Theories differ in how they conceive of and define EXH, how many EXH operators one needs, and in which positions they must occur in order to generate FREE CHOICE. These differences will not be important for our purposes. What is relevant for us is that EXH gives rise to FREE CHOICE as an implicature.

To illustrate how this approach works, let us return to the positive case in (1a), repeated schematically below in (12).

(12) $\Diamond(B \vee C) = \Diamond B \vee \Diamond C$ LITERAL MEANING

The literal meaning of (12) does not entail FREE CHOICE; in fact it is simply equivalent to $\Diamond B \vee \Diamond C$ (there is at least one between the boat and the car that Angie is allowed to buy).

The weak meaning in (12), when negated, immediately gives rise to the strong DOUBLE PROHIBITION reading: (13) entails $\neg \Diamond B \wedge \neg \Diamond C$ (Angie is not allowed to buy the boat and she isn't allowed to buy the car).

(13) $\neg[\Diamond(B \vee C)] = \neg \Diamond B \wedge \neg \Diamond C$ DOUBLE PROHIBITION

To generate FREE CHOICE, EXH is added to the positive sentence, generating the desired inference as an implicature.

(14) $\text{EXH}[\Diamond(B \vee C)] = \Diamond B \wedge \Diamond C$ FREE CHOICE

Moreover, the principle in (11) prevents, or makes it very hard for EXH to appear in the scope of negation, which would otherwise give rise to a NEGATED FREE CHOICE reading, (15), rather than the stronger DOUBLE PROHIBITION above in (13).

(15) $^{*}\neg[\text{EXH}[\Diamond(B \vee C)]] = \neg \Diamond B \vee \neg \Diamond C$ NEGATED FREE CHOICE

The derivations of the different meanings are schematized in Table 1.

Finally, the implicature approach can also account for the absence of FREE CHOICE and DOUBLE PROHIBITION in cases like (9) and (10). The latter can simply be a case in which EXH is not added to (12), giving rise to the weaker literal meaning that does not entail FREE CHOICE. This meaning is compatible with the continuation, where the speaker is explicit about being ignorant as to which of the boat and the car Angie is allowed to buy. (9) would be a case in which EXH does appear under negation (in the formulation of (11), EXH is allowed to appear under negation even

	LITERAL MEANING	IMPLICATURE	RESULT
POS	$\Diamond(B \vee C) = \Diamond B \vee \Diamond C$	$\text{EXH}(\Diamond(B \vee C)) = \Diamond B \wedge \Diamond C$	$\Diamond B \wedge \Diamond C$
NEG	$\neg\Diamond(B \vee C) = \neg\Diamond B \wedge \neg\Diamond C$	$^{*}\neg(\text{EXH}(\Diamond(B \vee C))) = \neg(\Diamond B \wedge \Diamond C)$	$\neg\Diamond B \wedge \neg\Diamond C$

Table 1 Derivation of the FREE CHOICE-DOUBLE PROHIBITION pattern under the implicature account.

if it leads to weakening, when this is 'forced' by the context). In this case, the potentially contradictory continuation would force EXH to appear under negation, giving rise to the NEGATED FREE CHOICE meaning in (15), which is compatible with the rest of the sentence and can therefore account for the felicity of (9).

In sum, given standard assumptions about the meanings of modals and disjunction, along with a theory of implicatures and a principle that regulates their distribution, the implicature approach can account for the FREE CHOICE-DOUBLE PROHIBITION pattern as well as cases in which these readings are absent. We turn now to the alternative, non-implicature approach.

2.2.2 A non-implicature approach: homogeneity

The implicature approach is not the only theoretical option available for explaining the FREE CHOICE-DOUBLE PROHIBITION pattern (see Zimmerman 2000, among others). Recently, a variety of alternative accounts have been proposed (e.g., Aloni 2007, Fusco 2015, Starr 2016, Willer 2017, Rothschild & Yablo 2018, Goldstein 2019, Aloni 2022). For concreteness, we focus on the HOMOGENEITY account in Goldstein (2019).[3] As far as we can see, however, the main points below apply to most of the non-implicature accounts; see also Aloni (2022) for discussion of how our results can be accounted for under her approach.

The HOMOGENEITY account is based on four ingredients: (i) a strong meaning for sentences like (1a), which directly asserts FREE CHOICE, (ii) a homogeneity presupposition requiring that either all alternatives are possible or none of them are, (iii) an operator that has the effect of cancelling the strong FREE CHOICE meaning, and (iv) another operator that can suspend DOUBLE PROHIBITION.

To illustrate, consider the positive case in (16a). Given (i), FREE CHOICE is directly entailed.[4] In addition, (16a) also presupposes HOMOGENEITY, as in (16b) (which in this case is entailed by the asserted FREE CHOICE meaning).

3 More specifically, we focus on his first account, based on alternative semantics.

4 See Goldstein (2019) for two possible ways of implementing this compositionally, by tweaking the standard meaning of modals or that of disjunction.

(16) a. $\Diamond(B \vee C) = \Diamond B \wedge \Diamond C$ FREE CHOICE
b. $\Diamond B \leftrightarrow \Diamond C$ HOMOGENEITY

While this approach directly captures FREE CHOICE by encoding it in the meaning of the positive sentence, it fails to immediately account for DOUBLE PROHIBITION. This is because in the negative case in (17a), the asserted meaning is now simply the negation of FREE CHOICE. However, the latter, in combination with the HOMOGENEITY presupposition in (17b) (which projects through negation), gives rise to the desired DOUBLE PROHIBITION reading in (17c).

(17) a. $\neg \Diamond (B \vee C) = \neg(\Diamond B \wedge \Diamond C)$ NEGATED FREE CHOICE
b. $\Diamond B \leftrightarrow \Diamond C$ HOMOGENEITY
c. $\neg \Diamond B \wedge \neg \Diamond C$ DOUBLE PROHIBITION

Finally, the HOMOGENEITY approach can capture examples like (9) and (10) by introducing two additional mechanisms. The first is an operator '!' that has the effect of cancelling FREE CHOICE when merged below the modal.[5]

(18) $\Diamond(!(B \vee C)) = \Diamond B \vee \Diamond C$

The second is a local accommodation operator '$\mathcal{A}$' (invoked independently in the presupposition literature, see Beaver (2001) and Fox (2012), among others), which, when merged below negation, makes the homogeneity presupposition an entailment and prevents it from projecting. The resulting meaning is the weak one in (19).

(19) $\neg[\mathcal{A}[\Diamond(B \vee C)]] =$
$\neg[(\Diamond B \wedge \Diamond C) \wedge (\Diamond B \leftrightarrow \Diamond C)] =$
$\neg[(\Diamond B \wedge \Diamond C)] = \neg \Diamond B \vee \neg \Diamond C$ NEGATED FREE CHOICE

In sum, combining the asserted meaning and the homogeneity presupposition, the homogeneity approach can capture the basic pattern, as well as the suspension of the FREE CHOICE and DOUBLE PROHIBITION readings (through the use of two additional operators). Table 2 provides a schematic illustration of the derivations of the different meanings under this approach.

2.2.3 Summary

The two theoretical approaches we have described can account for the basic FREE CHOICE and DOUBLE PROHIBITION readings, as well as cases in which these readings

5 The operator ! essentially has the effect of double negation; we again refer the reader to Goldstein (2019) for details.

	ASSERTED MEANING	PRESUPPOSED MEANING	RESULT
POSITIVE	$\Diamond(B \vee C) = \Diamond B \wedge \Diamond C$	$\Diamond B \leftrightarrow \Diamond C$	$\Diamond B \wedge \Diamond C$
NEGATIVE	$\neg\Diamond(B \vee C) = \neg(\Diamond B \wedge \Diamond C)$	$\Diamond B \leftrightarrow \Diamond C$	$\neg\Diamond B \wedge \neg\Diamond C$

Table 2 Derivation of the FREE CHOICE-DOUBLE PROHIBITION pattern under the HOMOGENEITY account.

appear to be absent. On the empirical end of things, a variety of more complex data points have been discussed in the literature, including free choice in the scope of universal and negative existential quantifiers (Chemla 2009, Van Tiel 2012, Bar-Lev & Fox 2017, Bar-Lev 2018, Bar-Lev & Fox 2020), modified numerals (Gotzner et al. 2017, to appear), free choice beyond disjunction (Chierchia 2013, Marty et al. 2021), the interaction between free choice and presuppositions (Romoli & Santorio 2019, Marty & Romoli 2019, Del Pinal et al. 2022), as well as the processing of free choice (Chemla & Bott 2014) and its acquisition in young children (Tieu et al. 2016). The jury is still out on which of the two approaches sketched above can best account for the observed empirical landscape.[6] In what follows, we will focus on diverging predictions of the two approaches.

Before moving on to our experimental predictions for the two approaches, we would like to briefly mention two other theoretical possibilities that have been explored in the literature. The first is the account in Barker (2010), which encodes FREE CHOICE in the semantics while deriving DOUBLE PROHIBITION as an implicature. This approach makes the opposite prediction of the standard implicature approach: in a context in which Angie is only allowed to buy the boat, a sentence like (1a) is predicted to be false, while its negative counterpart in (2a) is predicted to be an implicature violation. In other words, this approach predicts an asymmetry between the positive and negative free choice cases, but in the opposite direction of the corresponding disjunction cases. As we will see below, our results are also challenging for this kind of implicature approach. The second is a recent version of the implicature account, proposed by Del Pinal et al. (2022), in which the homogeneity presupposition is obtained through an implicature operator. This approach is of interest here because its predictions essentially align with those of the homogeneity approach in the free choice and double prohibition cases. As we will discuss below, this aspect of the approach is in line with our results. However,

6 While the non-implicature approach can more straightforwardly account for certain observed differences between free choice and other implicatures, one might argue it is more stipulative in nature. In particular, while the implicature approach does not need any extra assumptions about the meanings of modals and disjunction, the non-implicature approach needs to tweak these meanings in particular ways (see Bar-Lev 2018, Romoli & Santorio 2019, and Aloni 2022 for discussion).

given that this approach treats free choice and implicatures in the same way, it makes the same predictions for the implicature conditions, which is rather *not* in line with the clear differences that we observe between the free choice and implicature conditions, across the four experiments. We will discuss this further below.

2.3 Predictions

The two approaches we have described make similar predictions for the basic FREE CHOICE-DOUBLE PROHIBITION pattern, as well as for a variety of related data. There is one prediction where they diverge, however, which to our knowledge has been untested (although similar discussions exist in the context of plurals, see Križ 2015, Križ & Chemla 2015, Tieu et al. 2017a, Renans et al. 2018). This divergent prediction has to do with the status of the basic positive and negative cases. In particular, recall that under the implicature approach, FREE CHOICE arises as an implicature, while DOUBLE PROHIBITION is simply a part of the literal meaning.

(1a) Angie is allowed to buy the boat or the car.
↝ *Angie can choose between the two* IMPLICATURE

(2a) Angie is not allowed to buy the boat or the car.
↝ *Angie is not allowed to buy either one* LITERAL MEANING

Under the homogeneity approach, on the other hand, FREE CHOICE is part of the literal meaning, while DOUBLE PROHIBITION arises via the homogeneity presupposition. Crucially, under this approach, both the positive and negative cases are associated with the same homogeneity presupposition.

(1a) Angie is allowed to buy the boat or the car.
↝ *Angie can buy one iff she can buy the other* HOMOGENEITY

(2a) Angie is not allowed to buy the boat or the car.
↝ *Angie can buy one iff she can buy the other* HOMOGENEITY

The differences in the status of the positive and negative sentences can be brought out in a context in which only one of the disjuncts is allowed (e.g., Angie is only allowed to buy the boat). In this context, the homogeneity account predicts both the positive and negative cases to be undefined, as their presupposition is not satisfied. The implicature account, on the other hand, predicts a difference in status across the two polarities: it predicts the positive case to be a literally true sentence, but with a false implicature, while it predicts the negative case to be plainly false. The predictions are summarized in Table 3.

To sharpen the intuitions, consider the comparison with the corresponding

	Implicature	Homogeneity
Positive	IMPLICATURE VIOLATION	PRESUPPOSITION FAILURE
Negative	FALSITY	PRESUPPOSITION FAILURE

Table 3 Predictions of the implicature and homogeneity approaches for sentences like (1a) and (2a), in contexts in which only one of the disjuncts is allowed.

	Disjunction	Free choice (IMP)	Free choice (HOM)
Positive	IMP VIOLATION	IMP VIOLATION	PRESUPPOSITION FAILURE
Negative	FALSITY	FALSITY	PRESUPPOSITION FAILURE

Table 4 Predictions of the implicature and homogeneity approaches for sentences like (1a)/(2a) and (20a)/(21a), in contexts in which only one of the disjuncts is allowed (for free choice) and both of the disjuncts are true (for simple disjunction).

simple disjunction case. (20a) gives rise to an exclusivity implicature suggesting that Angie did not buy both the boat and the car. This inference disappears under negation: (21a) does not suggest that Angie bought both the car and the boat or neither of them, which would be the negation of (20a) with its exclusivity implicature. In other words, in a context in which Angie bought both the boat and the car, (20a) is predicted to be true but with a false implicature, while (21a) is predicted to be plainly false.

(20) a. Angie bought the boat or the car.
b. ⇝ *Angie didn't buy both the boat and the car* IMPLICATURE

(21) a. Angie didn't buy the boat or the car.
b. ⇝ *Angie didn't buy either one* NEGATED LITERAL MEANING

We can state the predictions as follows: the implicature approach predicts a similar pattern for the pairs in (1a)/(2a) and (20a)/(21a), reflecting a false implicature in the positive cases and a false literal meaning in the negative cases. The homogeneity account, on the other hand, is compatible with a difference between the pairs, in that (1a) and (2a), unlike (20a) and (21a), are predicted to have the same status (they are both predicted to be undefined). These predictions can be recast as in Table 4.

Quantitatively, observing a statistical interaction between Inference Type (EXCLUSIVITY VS. FREE CHOICE) and Polarity (positive vs. negative) would present

a challenge for the implicature approach, but would be entirely in line with the homogeneity approach. Testing these predictions gives us a simple way to distinguish between the two theoretical approaches. Before turning to our experiments, we will briefly outline some previous studies that the experiments build upon.

3 Previous studies

In recent years, a variety of studies have made use of a ternary judgment task to investigate implicatures (Katsos & Bishop 2011, Tieu et al. 2017a, Renans et al. 2018), presuppositions (Abrusan & Szendroi 2013, Zehr 2014), plural definite descriptions (Tieu et al. 2019, Augurzky et al. 2023), donkey pronouns (Sun et al. 2019), and counterfactuals (Marty et al. 2020).

As a task designed to test scalar implicatures (Katsos & Bishop 2011), the idea is that the lowest valued judgment (e.g., the smallest reward) is reserved for false sentences, the highest valued judgment (e.g., the biggest reward) is reserved for true and felicitous sentences, and the intermediate judgment then maps to true but infelicitous sentences, e.g., sentences with a true literal meaning but a false implicature. This ternary judgment scale has also been used to test other phenomena, including homogeneity in plural definite descriptions. Tieu et al. (2019), building on work with adults in Križ & Chemla (2015), used Katsos and Bishop's ternary judgment task to investigate children's sensitivity to homogeneity failures associated with plural definite descriptions (e.g., sentences like *The trucks are blue* and *The trucks are not blue* in a context in which only two out of four trucks were blue). The adult controls in their study generally responded in the same way to the positive and negative homogeneity failures — predominantly selecting intermediate and minimal rewards.[7] In conceiving the present study then, we intended a broader category of sentences to map to the intermediate (non-minimal/non-maximal) response option, which could in principle include sentences with false implicatures as well as homogeneity and presupposition failures.[8]

7 Križ & Chemla (2015) presented one group of adult participants with the response options 'Completely true' and 'Not completely true', and another group with the response options 'Completely false' and 'Not completely false'. Their participants generally treated homogeneity failures as being 'Not completely true' and 'Not completely false' — indicative of a truth value gap.

8 Zehr (2014) also used a ternary judgment task to test presupposition failures such as *The match has stopped burning* to describe a picture in which the match was never lit. Using the response options 'Completely true', 'Completely false', and 'Neither', Zehr observed that participants opted for either the 'Neither' option or the 'Completely false' option, but never the 'Completely true' option (note the parallel with the homogeneity results reported in Tieu et al. (2019)). We assume that presupposition failures might thus map to intermediate or minimal rewards, on a ternary response scale such as the one used in our study.

Tieu et al. (2017a) made use of a ternary judgment task to investigate the comprehension of plural noun phrases, comparing the multiplicity inference of bare plurals to the exclusivity inference of disjunction. In their study, participants saw pictures of cartoon characters deciding what to buy at the store, and heard a puppet's guesses as to what would or wouldn't happen (the so-called 'prediction' mode of the judgment task; see Tieu et al. 2017b, among others). Participants then saw the outcome and had to decide whether to reward the puppet with a small, medium, or large strawberry.

In particular, on the disjunction trials, participants heard positive and negative sentences such as (22) and (23), in contexts in which both disjuncts turned out to be true; for example, Tiger ended up buying both the apple and the banana, leaving behind a third object.

(22) Tiger will buy the apple or the banana.

(23) Tiger will not buy the apple or the banana.

In the target context in which both disjuncts turn out to be true, the positive (22) ends up being a true sentence with a false implicature, while its negative counterpart in (23) is plainly false. Participants were therefore expected to select the intermediate reward for (22), and the minimal reward corresponding to plain falsity for (23).

The results indeed revealed a clear contrast between plain falsity and implicature violation. Participants favored the intermediate reward for the positive targets, and gave mostly minimal rewards for the negative targets. This is in line with an implicature approach to the exclusivity inference, with participants mapping the intermediate reward to implicature violation and the minimal reward to falsity. This result is important because it provides a baseline against which to compare other phenomena that have also been subjected to an implicature analysis, such as the multiplicity inferences of plural noun phrases (Tieu et al. 2017a). We will use the same logic in the present study, comparing regular implicatures to free choice inferences.

In the following sections, we describe our set of four experiments. Experiment 1 focused on the comparison between free choice disjunction and plain disjunction. To anticipate, the results provide suggestive evidence against the implicature approach, but we identify a possible confound associated with intermediate reward selections in the ternary judgment task: participants may have potentially used the intermediate reward as a way to respond charitably to the puppet's statements. In Experiment 2a, we compared free choice disjunction to the weak modal *is allowed to*, and in Experiment 2b we introduced training and control trials that would allow us to circumvent the charitability confound. In Experiment 3, we compared free

choice 'any' to the scalar implicature of the quantifier 'some', while continuing to include the same training and controls that would address the charitability confound. To anticipate, all four experiments revealed a significant interaction between inference type and polarity, counter to the predictions of the scalar implicature account of free choice.

4 Experiment 1

4.1 Methods

4.1.1 Participants

120 participants were recruited through Prolific and randomly assigned to the free choice (FC) (n=60) or disjunction (OR) condition (n=60). Participants were prescreened for native language (English), location (USA, UK), and a previous approval rate of at least 90%. Participants were paid at an hourly rate of 10 GBP/hour. (The study took on average 6m10s to complete.)

4.1.2 Procedure

The procedure was the same in all three experiments that we will describe. All three experiments involved a ternary judgment task (Katsos & Bishop 2011) and were implemented using the Qualtrics platform. Participants were directed from Prolific to the experiment on the Qualtrics site.

In Experiment 1, participants were given a back story about characters who had each gone to the store. In each case, a puppet named Raffie would make a guess about what the character was allowed/not allowed to buy (FC condition) or about what the character had bought/hadn't bought (OR condition). Participants then had to decide upon seeing the pictured outcome how right the puppet had been, by selecting a reward for the puppet. Following the method in Katsos & Bishop (2011), we provided participants with three response options: a small strawberry (described in the instructions as corresponding to 'totally wrong' guesses), a large strawberry (for 'totally right' guesses), and a medium-sized strawberry (for when the puppet's guess was 'in between, not totally right but not totally wrong'). Participants indicated their response by clicking on the button that had the desired reward on it (Figure 1).

4.1.3 Materials

Each of the trials involved a display containing three objects. In the FC condition, we used a green circle around an object to indicate that the character was 'allowed'

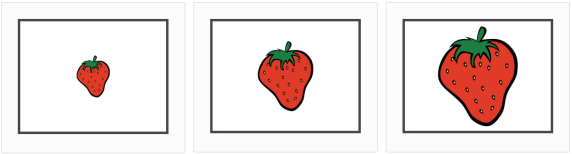

Figure 1 Response buttons for the ternary judgment task.

to buy the object, and a red circle with a line through it to indicate an object that a character was 'not allowed' to buy. Similarly, in the OR condition, objects that had been purchased were circled in green, and objects that had not been bought had a red circle with a line through it (see Figure 2 for examples). The type of inference (FC vs. OR) was a between-subjects factor, so participants did not have to change how they interpreted the green and red circles – this stayed the same throughout the experiment.

In the FC condition, the critical positive and negative target sentences (e.g., (1a) and (2a), repeated below) were presented in contexts that falsified the free choice inference, such as Figure 2 [left], in which only one of the disjuncts is 'allowed' (i.e. Angie is only allowed to buy the boat).

(1a) Angie is allowed to buy the boat or the car.

(2a) Angie is not allowed to buy the boat or the car.

In the OR condition, the positive and negative target sentences (e.g., (24) and (25)) were presented as guesses, to make the disjunction felicitous. The eventual pictured outcome against which the guesses were judged (e.g., Figure 2) [right] were incompatible with the exclusivity inference (i.e. Angie ended up buying both the boat and the car).

(24) Angie will buy the boat or the car.

(25) Angie will not buy the boat or the car.

In addition to the target items, participants received clearly true and clearly false (positive and negative) control items; see Figure 3 for the FC controls and Figure 4 for the OR controls.

Finally, alongside the clearly true and clearly false FC/OR controls, we also included four controls that we will refer to as 'partial truth' controls. These involved cases where two objects were mentioned using a *conjunction*, like (26) and (27), but the pictured context only had a green circle around one of the mentioned objects (e.g., Nina was only allowed to buy the peach, or Nina only purchased the peach).

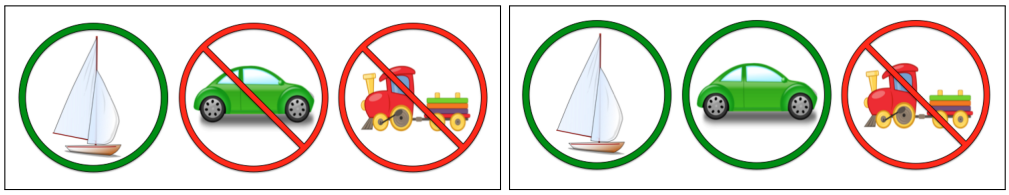

Figure 2 Example visual stimuli for FC and OR targets: the image on the left would be paired with the positive and negative FC targets in (1a) (*Angie is allowed to buy the boat or the car*) and (2a) (*Angie is not allowed to buy the boat or the car*); the image on the right would be paired with the positive and negative OR targets in (24) (*Angie will buy the boat or the car*) and (25) (*Angie will not buy the boat or the car*). (Actual items varied in the character's name and the pictured objects.)

(26) Nina is allowed to buy the peach and the carrot.

(27) Nina will buy the peach and the carrot

Given the use of conjunction, these sentences should be uncontroversially false in the pictured contexts. Pilot experiments had revealed that some participants might be charitable in responding to the puppet's guesses, such that if the puppet turned out to be right about at least one of the mentioned objects, they would choose the intermediate reward. We reasoned that including these 'partial truth' controls would give us some measure of this tendency.

In total, participants received 2 training items, followed by a fully randomized sequence of the 8 targets (4 positive, 4 negative), 8 true/false controls (2 true and 2 false positive controls, 2 true and 2 false negative ones), and 4 partial truth controls, for a total of 20 experimental trials.[9]

4.2 Results and discussion

As can be seen in Figure 5, participants primarily gave the intermediate reward in response to both positive and negative FC targets, while they gave an asymmetric pattern of responses to the positive and negative OR targets. We fitted a mixed effects cumulative link model to responses to the targets with Inference Type, Polarity, and their interaction as fixed effects, and random by-participant slopes for Polarity. Model comparisons between the maximal model and those

9 Note that every pictured context contained three objects, to avoid any potential infelicity associated with the use of a disjunctive statement to describe a context in which there are only two relevant objects (see Skordos et al. 2020 and Huang & Crain (2019) for relevant discussion).

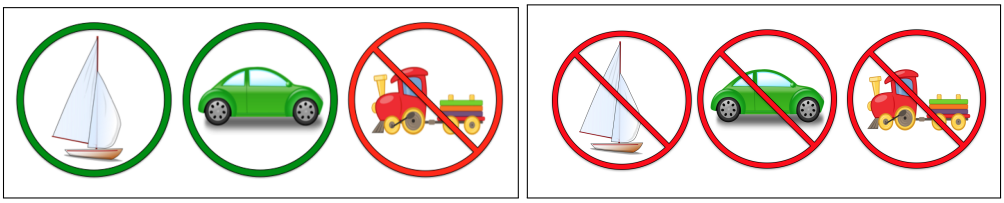

Figure 3 Example visual stimuli for the clearly true and clearly false FC controls. The positive sentence *Angie is allowed to buy the boat or the car* would be paired with the image on the left to create a true control and with the image on the right to create a false control. The negative sentence *Angie is not allowed to buy the boat or the car* would be paired with the image on the right to create a true control and with the image on the left to create a false control. (Actual items varied in the character's name and the pictured objects.)

without each of the fixed effects revealed a significant effect of Inference Type ($\chi^2(1) = 20, p < .001$), a significant effect of Polarity ($\chi^2(1) = 119, p < .001$), and a significant interaction between Inference Type and Polarity ($\chi^2(1) = 92, p < .001$), with participants distinguishing between the two polarities more so for OR than for FC.

The interaction between Inference Type and Polarity challenges the prediction of the implicature approach, namely that the free choice inference and the exclusivity implicature should behave similarly across polarities. There are two issues, however, which may confound our interpretation of the results, and will motivate our move to Experiments 2 and 3.

The first issue has to do with participants' behaviour on the controls trials for the plain disjunction OR. Figure 6 displays the results for the FC and OR controls. Participants' responses to the FC controls were as expected, with mostly maximal rewards assigned to the clearly true controls and minimal rewards assigned to the clearly false controls. In the disjunction condition, however, the positive true and negative false controls elicited a relatively large proportion of intermediate responses. These intermediate responses are surprising: in the positive case, the context was, for example, one in which Angie bought only the boat, thus (24) should have been judged as clearly true, and in the negative case, the same context was paired with (25), which therefore should have been judged as clearly false.

We suspect that the observed intermediate responses to what should have been clearly true disjunction controls may be related to potential residual effects of the ignorance inference of disjunction (despite our best efforts to circumvent the issue

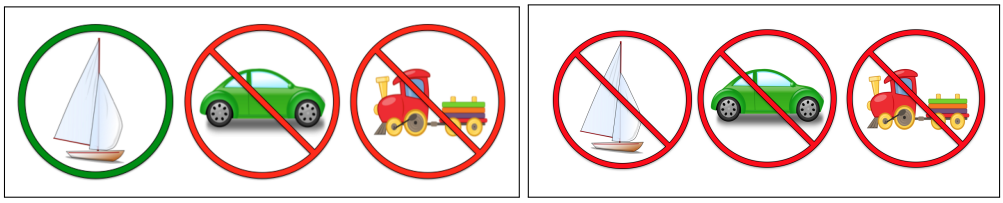

Figure 4 Example visual stimuli for the clearly true and clearly false OR controls. The positive sentence *Angie will buy the boat or the car* would be paired with the image on the left to create a true control and with the image on the right to create a false control. The negative sentence *Angie will not buy the boat or the car* would be paired with the image on the right to create a true control and with the image on the left to create a false control. (Actual items varied in the character's name and the pictured objects.)

by presenting the sentences in a predictive mode rather than in description mode). As for the intermediate responses to the negative controls, we speculate that they might be related to participants' preferences regarding the scope of disjunction relative to negation. Although the intended interpretation of negation scoping over disjunction should have led to a clear rejection of the control sentences, participants may have been able to access the inverse scope interpretation (which turned out to be true); participants sensitive to this ambiguity of the sentences may have thus assigned an intermediate reward to reflect the conflict between the false surface scope interpretation and the true inverse scope interpretation. We return to this hypothesis as potentially also applying to the free choice targets in Section 9.3.2 below. For now, we take this issue as a potential confound, which motivates the change in the scalar implicature baseline in the following experiments.

The second, more pressing issue, has to do with the partial truth controls we included in the experiment, in which conjunctive statements were presented in contexts that verified only one of the two conjuncts, thus rendering the sentences unambiguously false. We observed for these trials that the intermediate reward was nevertheless selected 86.3% percent of the time in the FC condition and 85.4% of the time in the OR condition. This suggests that participants may indeed have been adopting a kind of charitable response strategy: if the puppet turned out to be right about at least one of the mentioned objects, participants tended to select the intermediate reward. Importantly, on the free choice targets, one of the mentioned objects did end up being 'allowed' and one 'disallowed'. The possibility that participants could choose the intermediate reward because the puppet turned

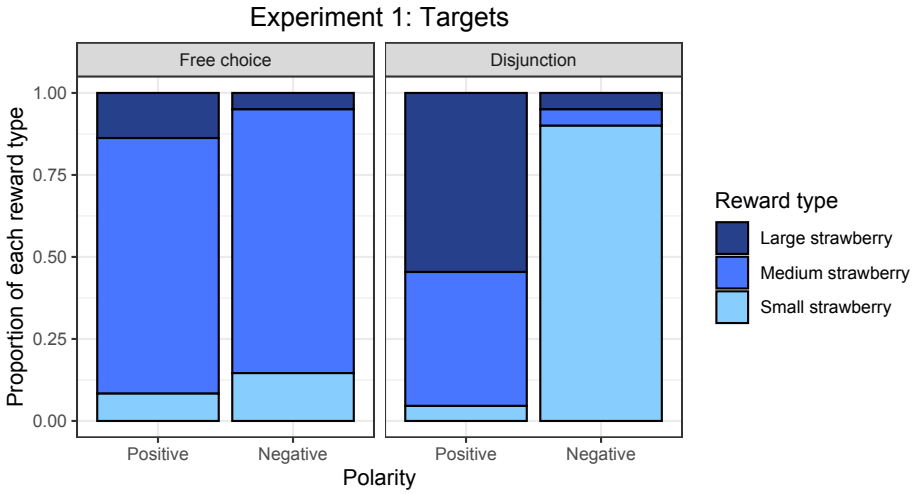

Figure 5 Experiment 1: Proportion of each reward type selected in response to free choice disjunction and plain disjunction targets.

out to be right about one of the mentioned objects thus weakens our ability to conclude that the intermediate responses to the free choice targets actually reflect undefinedness of the target sentences.

In the next three experiments, we will address each of these issues: in Experiment 2a we move away from plain disjunction to another inference that has been relatively uncontroversially treated as a scalar implicature, namely the inference associated with the weak modal *is allowed to*. The implicature of the modal allows us to avoid potential issues related to ignorance inferences of the disjunction. In Experiment 2b, we additionally include training trials to discourage participants from using the intermediate reward as a way of charitably rewarding 'partially true' statements, which we will see significantly decreases participants' use of this response strategy. In Experiment 3 we further attempt to circumvent the issue of charitability, by moving to the free choice quantifier 'any', with corresponding displays of nine, rather than three objects.

5 Experiment 2a

In Experiment 2a, we compared the free choice disjunction to the deontic modal *is allowed to* without the disjunction, to investigate potential differences in the behaviour of the free choice inference and the *not required to* inference of the modal verb, with respect to polarity. Alongside the FC condition then, we also tested

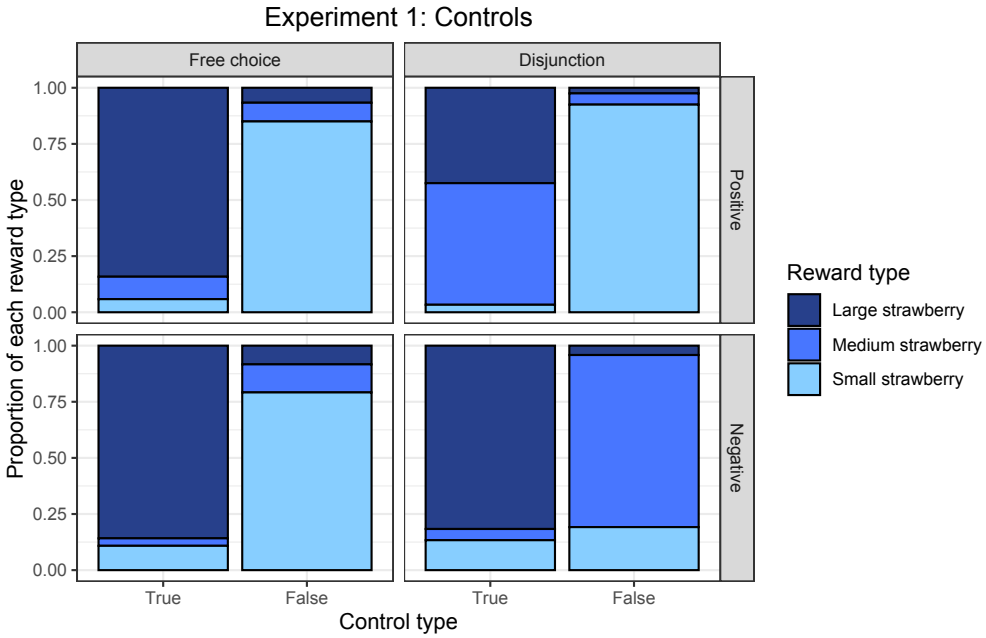

Figure 6 Experiment 1: Proportion of each reward type selected in response to clearly true and clearly false free choice and disjunction control items.

positive and negative deontic modal sentences like (28a) and (29a).

(28) a. Angie is allowed to buy the boat.
b. ⇝ *Angie is not required to buy the boat* IMPLICATURE

(29) a. Angie is not allowed to buy the boat.
b. ⇝ *Angie can't buy the boat* NEGATED LITERAL MEANING

As before, the implicature approach to free choice would predict no differences between the behaviour of the two inferences with respect to negation.

5.1 Methods

5.1.1 Participants

120 participants were recruited through Prolific and randomly assigned to the free choice (FC) (n=60) or modal condition (n=60). Participants were pre-screened for native language (English), location (USA, UK), and a previous approval rate of at least 90%. Using Prolific's pre-screening function, we also excluded people who had

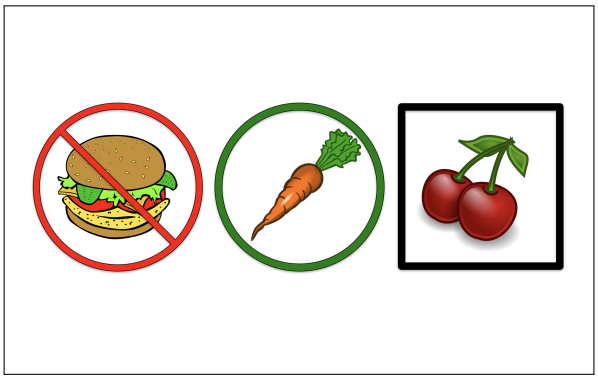

Figure 7 The instructions for Experiment 2 included this picture containing the different symbols that would be used in the experiment, along with this explanation: *Here is how we'll represent the rules. Take a look at the picture below. The red circle with the line through the hamburger means that the character is not allowed to buy the hamburger. The green circle around the carrot means that the character is allowed to buy the carrot. And the black square around the cherries means that the character has to buy the cherries.*

completed Experiment 1 from participating in Experiment 2. Participants were paid at an hourly rate of 10 GBP/hour. (The study took on average 5m33s to complete.)

5.1.2 Procedure

The task was the same ternary judgment task with three response options, as in Experiment 1, with the puppet's guess presented in text prior to the pictured outcome being displayed on the page.

Because we were testing the inference of the deontic modal, the three-object displays had to be able to represent not just possibility (green circle) and impossibility (red circle with a line through it), but also obligation. To do this, we introduced a black box that could outline certain objects, which participants were told meant that the character had to buy the relevant object. The relevant part of the instructions, along with the accompanying image, is provided in Figure 7.

5.1.3 Materials

The FC condition was a replication of the FC condition from Experiment 1. Participants saw all the same items that were tested in Experiment 1: 4 positive FC targets,

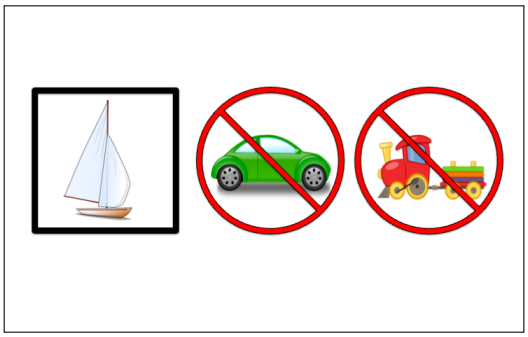

Figure 8 Example visual stimulus for the MODAL targets. This image would be paired with the positive and negative MODAL target sentences in (28a) (*Angie is allowed to buy the boat*) and (29a) (*Angie is not allowed to buy the boat*).

4 negative FC targets in contexts that falsified the free choice inference, 4 positive FC controls (2 clearly true, 2 clearly false), 4 negative FC controls (2 clearly true, 2 clearly false), and 4 clearly false conjunction controls, for a total of 20 experimental items, presented in randomized order.

The MODAL condition had the same structure with a total of 20 experimental items, but was adapted to the modal 'is allowed to'. Participants had to judge 4 positive MODAL targets such as (28a) and 4 negative MODAL targets such as (29a), presented in contexts like Figure 8, which falsified the *not required to* implicature of the deontic modal.

In addition to the target items, participants received four clearly true and four clearly false (positive and negative) control items, as illustrated in Figure 9. Finally, alongside the clearly true and clearly false modal controls, we also included four 'partial truth' conjunction controls, as in the FC condition (and as in Experiment 1).

5.2 Results

As can be seen in Figure 10, participants primarily gave the intermediate reward in response to both positive and negative FC targets, while they gave an asymmetric pattern of responses to the positive and negative modal targets. We fitted a mixed effects cumulative link model to responses to the targets with Inference Type, Polarity, and their interaction as fixed effects, and random by-participant slopes for Polarity. Model comparisons between the maximal model and those without each of the fixed effects revealed a significant effect of Inference Type ($\chi^2(1) = 31, p < .001$), a significant effect of Polarity ($\chi^2(1) = 106, p < .001$), and a

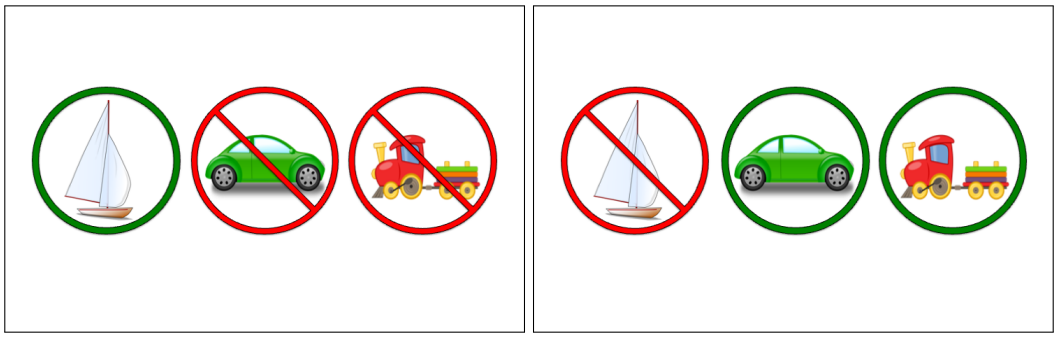

Figure 9 Example visual stimuli for the clearly true and clearly false MODAL controls. The positive sentence *Angie is allowed to buy the boat* would be paired with the image on the left to create a true control and with the image on the right to create a false control. The negative sentence *Angie isn't allowed to buy the boat* would be paired with the image on the right to create a true control and with the image on the left to create a false control. (Actual items varied in the character's name and the pictured objects.)

significant interaction between Inference Type and Polarity ($\chi^2(1) = 97, p < .001$), with participants showing a greater difference between polarities in the MODAL condition compared to the FC condition.

Responses were as expected for the clearly true and clearly false controls (Figure 11). The modal condition showed none of the unexpected effects observed for the plain disjunction controls in Experiment 1.

On the partial truth controls, however, we observed a similar effect as in Experiment 1, with participants selecting the intermediate reward for what should have been uncontroversially false conjunctive statements 90.8% of the time in the FC condition and 83.3% of the time in the MODAL condition. This suggests that in this experiment too, participants may have been adopting a charitable response strategy, whereby they were inclined to give an intermediate reward when the puppet turned out to be right about at least one of the mentioned objects.[10]

10 We also considered excluding participants who consistently adopted the charitable response strategy, namely those who accepted at least three of the four false conjunctive controls; but this would have led to the exclusion of the majority of participants, leaving only 3/60 in the free choice condition and 7/60 in the modal condition (a 92% participant exclusion rate). This speaks to how strong the tendency was to use the intermediate reward charitably.

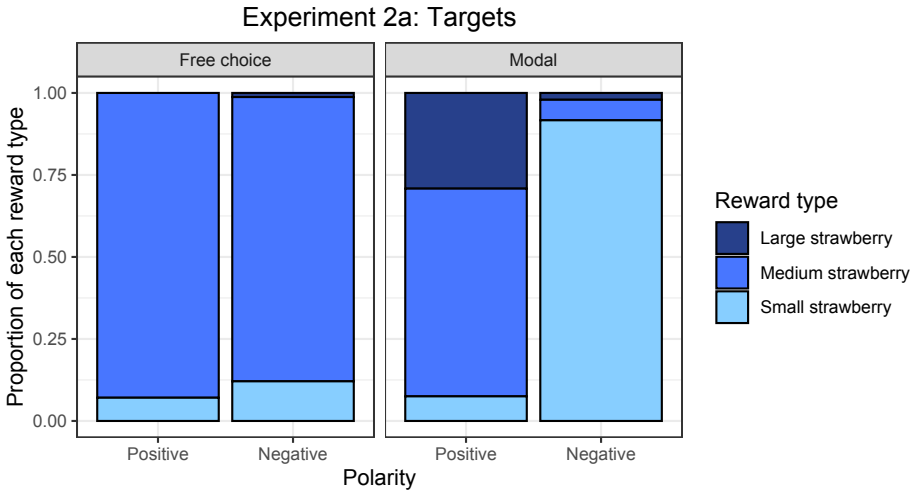

Figure 10 Experiment 2a: Proportion of each reward type selected in response to free choice and modal targets.

6 Experiment 2b: Extra training

Experiment 2b was the same as Experiment 2a with two differences. First, we added the word 'both' to the conjunctive controls to make them even clearer (e.g., *Kai is allowed to buy both the juice and the potato*). Second, and more substantively, we added four more training trials to the beginning of the experiment, the purpose of which was to encourage participants to choose minimal rewards for clearly false conjunctive statements. Before starting the test trials of the experiment, participants thus saw a total of seven training trials. Three were clearly true/clearly false trials and were the same as in Experiment 2a: two corresponded to clearly true simple statements (e.g., *Megan has to buy the hamburger* (modal condition) / *Megan is allowed to buy the hamburger* (FC condition)), while one was a clearly false simple statement (e.g., *Terry is not allowed to buy the cherries* in a context where Terry *was* allowed to buy the cherries). The four new conjunctive partial truth controls involved a conjunctive statement such as *Lulu is allowed to buy both the crown and the book* in a context in which Lulu was only allowed to buy the crown. As before, accompanying each unambiguous training trial was a directive about which of the three rewards participants should select, and why. For the partial truth training trials, the directives were of the form: *Was Raffie right? In this case, you should give Raffie the smallest strawberry: Lulu WASN'T allowed to buy both!*

Besides the addition of the four training trials, the design for the test phase of the experiment was the same as in Experiment 2a. We compared the free

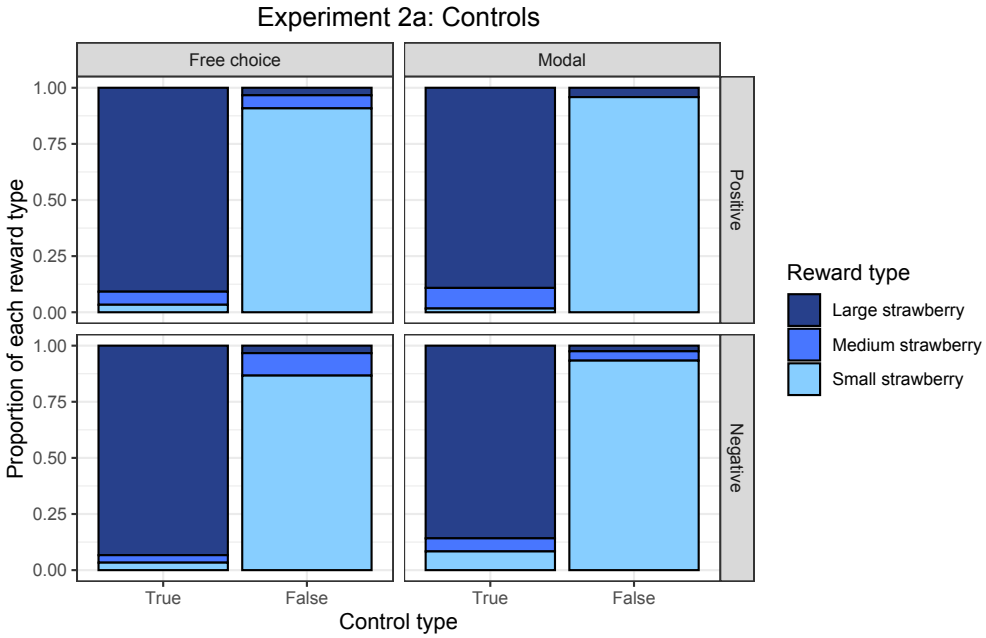

Figure 11 Experiment 2a: Proportion of each reward type selected in response to clearly true and clearly false free choice and modal controls.

choice disjunction to the deontic modal *is allowed to* without the disjunction, to investigate potential differences in the behaviour of the free choice inference and the *not required to* inference of the modal verb, with respect to polarity. As before, the implicature approach to free choice would predict no differences between the behavior of the two inferences with respect to negation.

6.1 Participants

A group of 182 participants were recruited through Prolific and randomly assigned to the free choice (FC) (n=91) or modal condition (n=91). Participants were pre-screened for native language (English), location (USA, UK), and a previous approval rate of 90-100%. Using the Prolific pre-screening function, we also excluded people who had previously completed Experiment 1 or Experiment 2a. Participants were paid at an hourly rate of 10 GBP/hour. (The study took on average 7m33s to complete.)

6.2 Procedure

The procedure was the same as in Experiment 2a.

6.3 Materials

The materials for Experiment 2b were the same as those for Experiment 2a, aside from the two differences described above: (i) the addition of four more training trials to discourage the use of the charitable 'partial truth' response strategy, and (ii) the addition of the word 'both' to the conjunctive partial truth controls (e.g., from *Nina is allowed to buy the peach and the carrot* to *Nina is allowed to buy both the peach and the carrot.* As before, participants saw a total of 20 experimental items, presented in randomized order.

6.4 Results

Before conducting our planned analyses, we first examined responses to the partial truth control conditions, which provided a measure of participants' tendency to rely on a charitable response strategy of rewarding partially true statements. Intermediate reward selections on the partial truth conjunctive controls corresponded to 34% of responses in the free choice condition, and 28% of responses in the modal condition (compared to 91% and 83% in Experiment 2a, in which we did not provide the extra partial truth training). The relatively lower rate of intermediate responses to the false conjunctive controls suggests that the additional training was effective at reducing participants' tendency to reward for partial truth.

We assumed that participants who, despite the training, failed to select the minimal reward on at least three of the four false conjunctive controls were those who might consistently use the intermediate reward charitably, and thus excluded these participants from our planned analyses. This exclusion criterion left a total of 56/91 participants in the free choice condition and 61/91 participants in the modal condition (a 36% participant exclusion rate, compared to what would have been a 92% participant exclusion rate in Experiment 2a, without training).

For the 56 participants in the free choice condition and 61 participants in the modal condition, performance on the unambiguous true/false controls was as expected (see Figure 12): minimal rewards for clearly false statements and maximal rewards for clearly true statements. Again, the modal condition showed none of the unexpected effects observed for the plain disjunction controls in Experiment 1.

Moving on to the target conditions, as can be seen in Figure 13, participants mostly gave minimal rewards in response to both positive and negative FC targets, while they gave an asymmetric pattern of responses to the positive and negative modal targets. We fit a mixed effects cumulative link model to responses to the

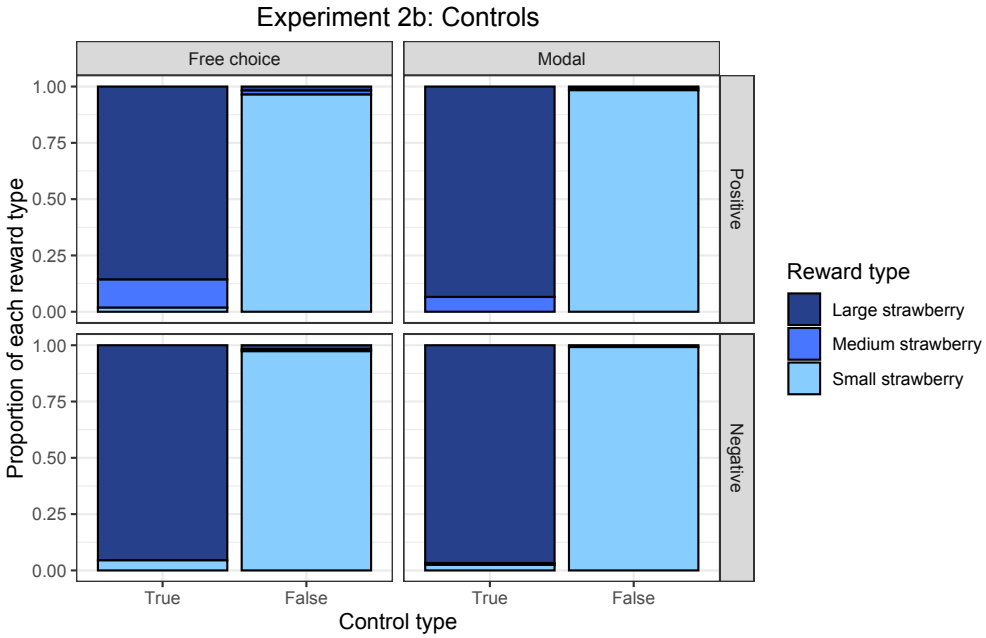

Figure 12 Experiment 2b: Proportion of each reward type selected in response to clearly true and clearly false free choice and modal controls.

targets with Inference Type, Polarity, and their interaction as fixed effects, and random by-participant intercepts. Model comparisons between the maximal model and those without each of the fixed effects revealed a significant effect of Inference Type ($\chi^2(1) = 8.8, p < .01$), a significant effect of Polarity ($\chi^2(1) = 492, p < .001$), and a significant interaction between Inference Type and Polarity ($\chi^2(1) = 315, p < .001$), with participants showing a greater difference between polarities in the MODAL condition compared to the FC condition.

On the whole, the results pattern with those of the two previous experiments, except that we can be more confident in our conclusions, having set aside the partial truth confound. After training participants away from selecting the intermediate reward for partially true statements, and moreover excluding participants who consistently responded charitably, we observe fewer intermediate rewards (notably in the free choice condition). Importantly, the main interaction still stands: participants differentiated between the positive and negative modal statements, but not between the positive and negative free choice statements.

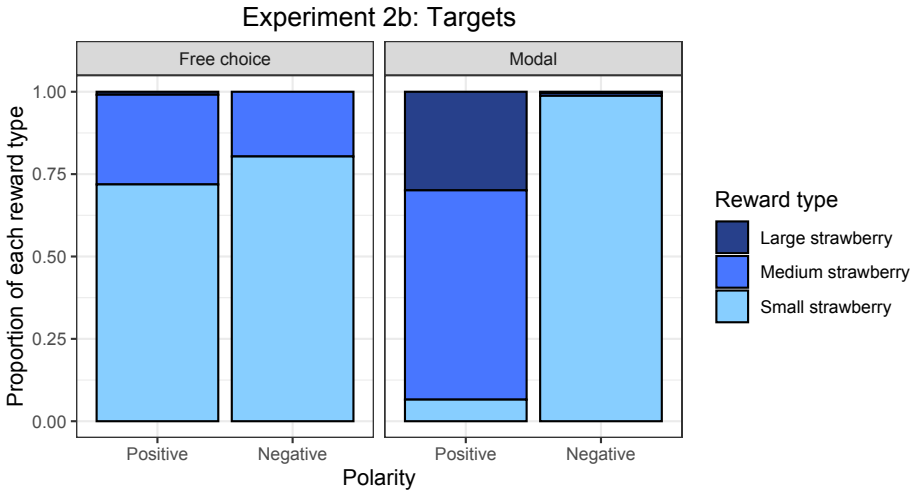

Figure 13 Experiment 2b: Proportion of each reward type selected in response to free choice and modal targets.

7 Interim discussion

Across Experiments 1 and 2a, we found that participants primarily selected the intermediate reward for both the positive (1a) and negative (2a) free choice sentences in the target contexts. In contrast, when presented with simple disjunctive sentences like (24) or a simple modal statement without disjunction like (28a) and their negative counterparts (25) and (29a), in the corresponding contexts, participants exhibited the asymmetric pattern of responses expected on the implicature approach: a preference for the intermediate reward when the (positive) sentence was logically true but had a false implicature, and the minimal reward when the (negative) sentence was plainly false. The parallel responses to (1a) and (2a), combined with the observed divergent responses to the equivalent disjunctive and simple modal sentences, posed a challenge for the implicature approach.

As mentioned, however, these results can also be explained as participants having chosen the intermediate reward in an attempt to be charitable to the puppet. That is, the puppet mentioned two things (let's say, the hamburger and the carrot) and she turned out to be right about one of them (the character did end up being allowed to buy the hamburger). So, while the sentence on its free choice meaning should not be compatible with the pictured context, there is a sense in which the puppet's guess was partially right, and this could underlie the observed intermediate responses.

And indeed, corroborating the potential presence of this charitable response

strategy, we observed a similar pattern of intermediate responses on our conjunction controls, where no implicature was involved at all (i.e. (26) or (27)), suggesting participants may simply have been responding charitably to the target items. This concern highlights a difficulty with the use of the ternary judgment task, insofar as there might be multiple explanations for why a participant opts for an intermediate reward. It seems particularly acute in this case, however, because we are dealing with sentences in which the puppet explicitly mentions two objects by name; when the puppet turns out to be right about one of them, it's easy to see the temptation to partially reward the puppet, and the ternary judgment scale offers just such a partial reward option.

With the inclusion of Experiment 2b, we arrive at the same overall finding, but with more confidence: Experiment 2b controlled for the charitable partial truth confound, by training participants away from using the intermediate reward to be charitable, and by excluding participants who persisted in using this response strategy. With a sizable group of participants remaining in each condition, we nevertheless continued to observe the main interaction effect, with participants failing to differentiate between polarities for free choice, but differentiating across polarities for the scalar implicature baseline.

In Experiment 3, we will move away from disjunctive free choice statements entirely, in a further attempt to mitigate the temptation to partially reward the puppet. Instead, we will investigate the phenomenon of free choice using the quantifier 'any', and compare it to the scalar implicature of 'some'. In particular, we will use sentences like (30a) and (31a) and compare them to (32a) and (33a). The logic is the same as outlined for the first three experiments: both approaches have been extended from free choice disjunction to free choice items like *any* and their predictions extend to this domain as well (see Aloni 2007 and Chierchia 2013, among others).

(30) a. Angie is allowed to buy any of the items.
b. ⇝ *Angie can freely choose amongst all the items* FREE CHOICE

(31) a. Angie is not allowed to buy any of the items.
b. ⇝ *Angie cannot buy any of the items* NEGATED LITERAL MEANING

(32) a. Angie bought some of the items.
b. ⇝ *Angie didn't buy all of the items* IMPLICATURE

(33) a. Angie didn't buy any of the items.
b. ⇝ *Angie didn't buy any of the items* NEGATED LITERAL MEANING

The hypothesis is that the larger domain of food items (we will use displays of 9 items) will encourage fewer partial truth responses. Moreover, to further strengthen the conclusions we can draw from the data, we continue to include the additional

partial truth/charitability training trials at the beginning of the experiment, as well as the conjunctive partial truth controls to quantify the potential use of a charitable response strategy, and ultimately to exclude participants who consistently exhibit this type of charitable response. A comparison of our target conditions with and without exclusion of the 'partial truth' responders will also shed light on the potential role of this strategy in explaining our results.

While both an implicature and a homogeneity account can be extended to free choice 'any', we should note that there are also differences between free choice 'any' and free choice disjunction that have been discussed in the literature. Most notably, free choice 'any' is considered to be subject to more stringent licensing constraints, and is associated with more of an 'obligatory' inference than free choice disjunction (Aloni 2007, Chierchia et al. 2012). The extent to which the results relating to 'any' in Experiment 3 will be able to speak to theories of free choice disjunction is not necessarily straightforward. Concretely, the predictions of the implicature approach for the positive condition are less clear than in the free choice disjunction case, as we do not know how an 'obligatory implicature' will map to ternary responses. Nonetheless, with these caveats in mind, we think that this comparison is a reasonable move and may be suggestive for accounts that treat both types of free choice inferences in a unified fashion.

8 Experiment 3

8.1 Methods

8.1.1 Participants

119 participants were recruited through Prolific and randomly assigned to the free choice 'any' (n=60) and 'some' conditions (n=59). Participants were pre-screened for native language (English) and location (USA, UK). Using Prolific's pre-screening function, we also excluded people who had completed any of the previous experiments. Participants were paid at an hourly rate of 10GBP/hour. (The study took on average 7m2s to complete.)

8.1.2 Procedure

The procedure was the same as in Experiments 1 and 2a/b, with participants being asked to judge a puppet's guesses against the pictured outcome, using the three-strawberry response options.

Figure 14 Sample image accompanying the positive and negative FC 'any' targets in (34a) (*Angie is allowed to buy any of the items*) and (35a) (*Angie is not allowed to buy any of the items*).

8.1.3 Materials

To adapt the free choice condition from free choice disjunction to free choice 'any', we modified the images to contain a 3x3 display of nine objects, instead of three. This would not only make the 'any' statements more natural, but would also allow us to compare 'any' to the quantifier 'some' under the same conditions.

In the free choice condition, the critical positive and negative target sentences (e.g., (30a) and (31a), repeated below as (34a) and (35a)) were presented in contexts that falsified the free choice inference, such as Figure 14, in which only four of the nine domain alternatives were actually 'allowed'.

(34) a. Angie is allowed to buy any of the items.
b. ⇝ *Angie can freely choose amongst the items* FREE CHOICE

(35) a. Angie is not allowed to buy any of the items.
b. ⇝ *Angie cannot buy any of the items* NEGATED LITERAL MEANING

Participants in the SOME condition saw positive 'some' and negative 'any' target

Figure 15 Sample image accompanying the positive and negative SOME targets in (36a) (*Angie bought some of the items*) and (37a) (*Angie didn't buy any of the items*).

sentences (e.g., (32a) and (33a), repeated below as (36a) and (37a)), accompanied by pictures which falsified the *not all* implicature of 'some' (e.g., Figure 15).

(36) a. Angie bought some of the items.
b. ⇝ *Angie didn't buy all of the items* IMPLICATURE

(37) a. Angie didn't buy any of the items.
b. ⇝ *Angie didn't buy any of the items* NEGATED LITERAL MEANING

In addition to the target items, participants received clearly true and clearly false (positive and negative) control items; see Figure 16 for the FC controls and Figure 17 for the SOME controls.

Finally, alongside the clearly true and clearly false FC/OR controls, we also included four 'partial truth' controls of the type that were included in Experiments 1 and 2a/b. To adapt the partial truth controls to Experiment 3, the items this time involved nine pictured objects, with four of them circled ('allowed'), such that that the accompanying 'every' sentence (e.g., (38)) should be clearly false — unless participants were adopting the charitable 'partial truth' response strategy.

Figure 16 Example visual stimuli for the clearly true and clearly false FC controls. The positive sentence *Angie is allowed to buy any of the items* would be paired with the image on the left to create a true control and with the image on the right to create a false control. The negative sentence *Angie isn't allowed to buy any of the items* would be paired with the image on the right to create a true control and with the image on the left to create a false control. (Actual items varied in the character's name and the pictured objects.)

(38) Nina is allowed to buy every item.

Additionally, in the training phase of the experiment, we followed the structure of Experiment 2b, which meant participants saw three clearly true/clearly false simple statements (e.g., *Megan is allowed to buy the hamburger*, *Terry is not allowed to buy the cherries*), and four partial truth training trials (e.g., *Lulu is allowed to buy every item* in a context in which she was only allowed to buy 4/9 of the pictured items). Similarly to Experiment 2b, the directives on the partial truth training controls were of the form: *Is Raffie right? In this case, you should give Raffie the smallest strawberry: Lulu ISN'T allowed to buy every item!*

In total, participants received seven training items, followed by a fully ran-

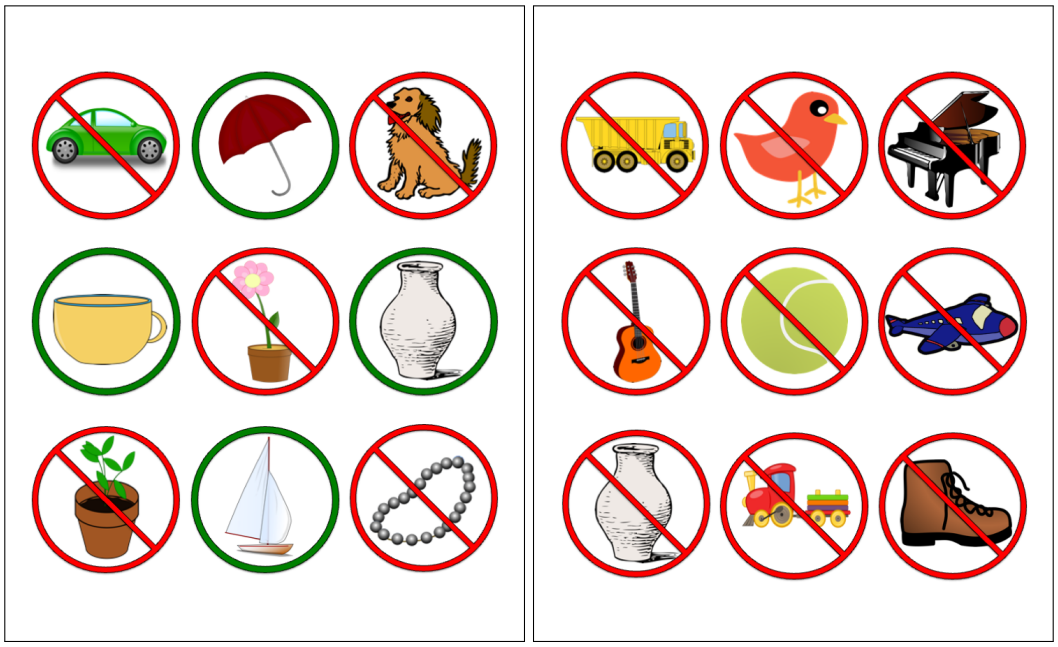

Figure 17 Example visual stimuli for the clearly true and clearly false SOME controls. The positive sentence *Angie is allowed to buy some of the items* would be paired with the image on the left to create a true control and with the image on the right to create a false control. The negative sentence *Angie isn't allowed to buy any of the items* would be paired with the image on the right to create a true control and with the image on the left to create a false control. (Actual items varied in the character's name and the pictured objects.)

domized sequence of the 8 targets (4 positive, 4 negative), 12 true/false controls (as before, we included four false controls, but this time had eight true controls instead of four, to balance out the expected number of *yes*-/*no*-responses), and 4 partial truth controls, for a total of 24 experimental trials.

8.2 Results

Let us first consider the 'every' partial truth controls, which were the equivalent of the conjunction controls in Experiments 1 and 2a/b. The first result worth noting is that the combination of the partial truth training and the move to free choice 'any' (which involved displays with more objects, and notably moving away

from explicitly pronouncing two disjuncts), appears to have effectively reduced participants' recourse to the charitable response strategy. In response to the clearly false 'every' controls, 13% of responses in the FC condition and 8% of responses in the SOME condition corresponded to the intermediate reward, compared to 83–91% intermediate reward selections for the partial truth controls in Experiments 1 and 2a.[11]

To be even more conservative, for our planned analysis, we first used performance on the partial truth controls to filter out participants who appeared to consistently adopt a charitable response strategy (in spite of the additional training). We eliminated from analysis participants who selected the intermediate reward on at least three out of the four partial truth controls. This criterion left 50/60 participants in the ANY condition and 53/59 participants in the SOME condition (a 13% participant exclusion rate, compared to the 36% exclusion rate in Experiment 2b (and a 39% exclusion rate in the pilot version of Experiment 3 that did not include the additional training).

For the remaining 50 participants in the ANY condition and 53 participants in the SOME condition, performance on the clearly true and clearly false controls was as expected, as seen in Figure 18: minimal rewards for clearly false statements and maximal rewards for clearly true statements.

As can be seen in Figure 19, participants primarily gave the minimal reward in response to both positive and negative FC targets, while they gave an asymmetric pattern of responses to the positive and negative SOME targets. We fit a mixed effects cumulative link model to responses to the targets with Inference Type, Polarity, and their interaction as fixed effects, and random by-participant slopes for Polarity. Model comparisons between the maximal model and those without each of the fixed effects revealed a significant effect of Inference Type ($\chi^2(1) = 7, p < .01$), a significant effect of Polarity ($\chi^2(1) = 36, p < .001$), and a marginally significant interaction between Inference Type and Polarity ($\chi^2(1) = 3.7, p = .055$), with participants showing a greater difference between polarities in the SOME condition compared to the FC condition.

8.3 Discussion

Experiment 3 replicates the finding of a difference between free choice and regular implicatures observed in Experiments 1 and 2a/b. As in Experiment 2b, we reduced

11 A pilot version of Experiment 3 without the training trials revealed 39% (FC) and 22% (SOME) intermediate reward selections for the false conjunctive controls. These rates are both lower than those observed in Experiments 1 and 2a, and higher than the ones we report above for Experiment 3 (with training), which suggests the move to 'any' and the presence of the additional training both played a role in reducing participants' recourse to the charitable response strategy.

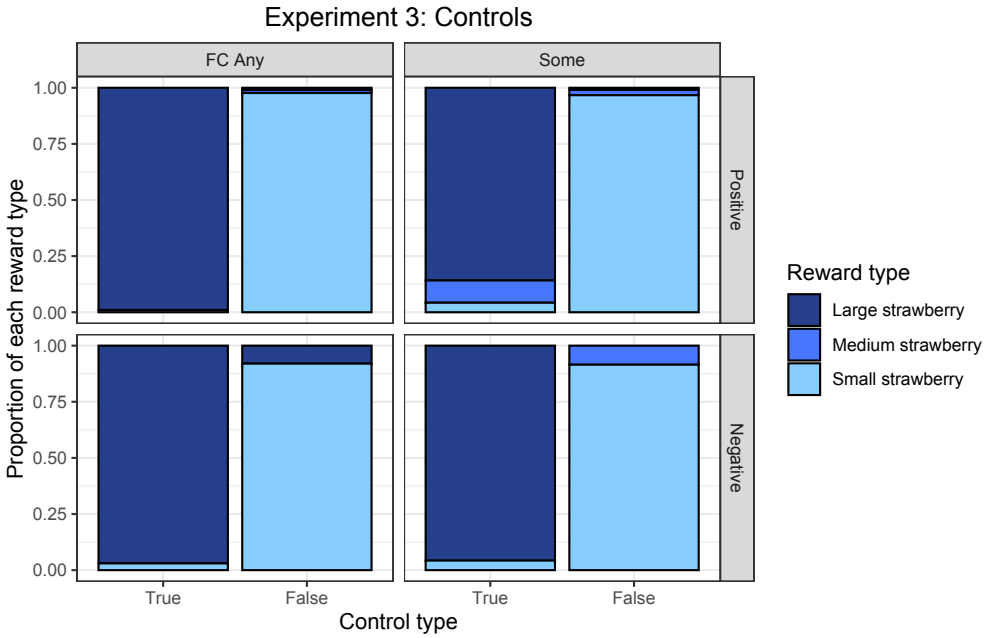

Figure 18 Experiment 3: Proportion of each reward type selected in response to free choice 'any' and 'some' controls, following exclusion of 'partial truth' responders.

the charitable response strategy through the addition of extra training trials, as well as partial truth controls to exclude participants who exhibited this strategy. As mentioned, while there are unified implicature and non-implicature accounts of free choice disjunction and free choice 'any', there are also differences between the two, in particular in terms of the 'obligatoriness' of the inferences and the constraints on their distribution. As a consequence, Experiment 3 may not test theories of free choice disjunction as straightforwardly as Experiments 1 and 2a/b do. Nevertheless, the comparison between the two phenomena and the very similar results we observe, once the partial truth strategy is controlled for in the same fashion, is rather suggestive for unified accounts.

9 General discussion

9.1 The challenge

Overall, the observed difference between the free choice and implicature targets, and in particular the observed interaction between Inference Type and Polarity, is

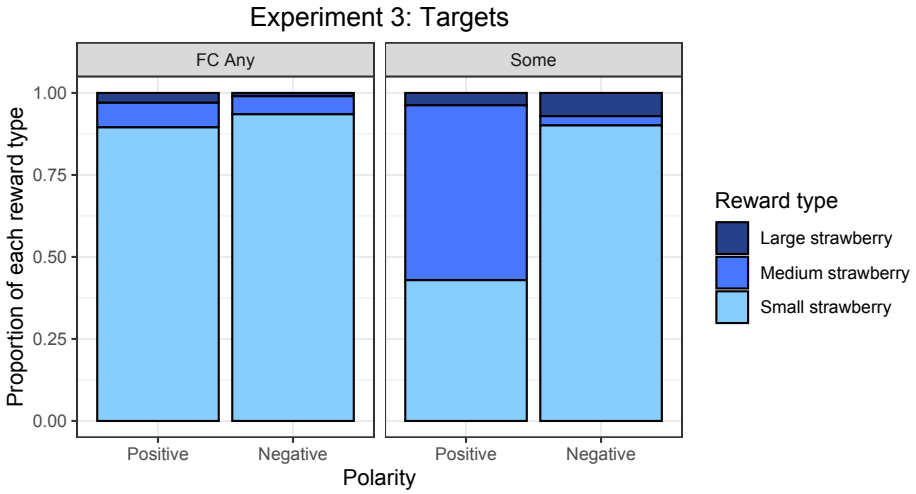

Figure 19 Experiment 3: Proportion of each reward type selected in response to free choice 'any' and 'some' targets, following exclusion of 'partial truth' responders.

challenging for the implicature account, which predicts a similar pattern across polarities for the two inference types. On the other hand, the results are straightforwardly in line with the homogeneity account, which predicts both positive and negative cases of free choice to be equally undefined in the given context, and is compatible with the observed interaction between Inference Type and Polarity.

Note that a response to the challenge against the implicature account cannot lie (entirely) in a 'scalar diversity' effect (van Tiel et al. 2016). That is, the observed interaction cannot be explained by appealing to a relative difference in the strength of the free choice and regular scalar inferences. This line of explanation would at most be able to account for the difference observed in the positive condition, but would be silent about the difference observed in the negative condition.

In addition, the nature of the *alternatives* involved in inference computation has been argued to play a role in observed differences between free choice and other inferences, in processing and acquisition (Chemla & Bott 2014, Tieu et al. 2016; see also Marty et al. 2021); however, it is not obvious to us how appealing to alternatives would be able to account for the present data.[12]

12 An anonymous reviewer suggests that a difference in the prunability of alternatives could be one possible route: if lexical alternatives (those involved in the exclusivity implicature baseline in Experiment 1) are easily pruned, but the individual disjunct alternatives involved in free choice disjunction are not prunable, this could underlie the observed 'optional' pattern of a mix of intermediate and

9.2 The ternary task and partial truths

As discussed, we set out to investigate the predictions of the two approaches by using a ternary judgment task, with the idea that the intermediate value could be straightforwardly interpreted as reflecting semantic undefinedness. Along the way, we discovered that many of our participants actually used this response option as part of what could be described as a charitable response strategy, to reward the puppet for being 'partially right'. We thus introduced conjunctive controls to be able to detect the presence of such a strategy, and in Experiment 2b and Experiment 3, provided additional training to discourage the use of such a response strategy. In Experiment 3, we additionally moved away from free choice disjunction, to free choice 'any', and used displays with a greater number of objects, to further discourage the temptation to 'partially reward' the puppet. In both Experiment 2b and Experiment 3, we indeed observed less evidence for use of the charitable strategy, and were able to exclude the relative minority of participants who persisted with the charitable response strategy. We can thus be reasonably confident that this strategy does not play a role in explaining the main findings.

Nonetheless, given the degree to which we observed 'partial truth' responses in Experiments 1 and 2a, we think that any future studies that use the ternary judgment task in the way we have used it here should crucially include the appropriate controls.

9.3 Two routes for the implicature approach

As mentioned, it is not clear how one might address the challenge to the implicature approach by appealing to the notion of scalar diversity, or to a difference in the alternatives involved in computing free choice and regular implicatures. Two possible alternative directions would be to reconsider the principle regulating the distribution of implicatures, or going back to a wide scope explanation combined with an hypothesis about how participants use the ternary task in the presence of ambiguity.

---

maximal rewards for the implicature baseline, but mostly intermediate ('pragmatic') responses for the free choice condition. While this could explain the pattern for the positive free choice items, it cannot explain the pattern in the negative cases. For the negative items, even with such assumptions about prunability, the scalar implicature approach predicts plain falsity, while participants tended to select the intermediate option. Having said this, we think the strongest evidence for our claims comes from Experiment 2b, where we better control for the partial truth strategy, and there the intermediate responses are substantially reduced. We think further exploration of this option as well as the relative prunability of different alternatives are interesting avenues for future research.

9.3.1 A novel constraint on implicature distribution

To illustrate, consider the standard principle in (11), which strongly disfavours the appearance of EXH in the scope of negation. As with our examples, the presence of EXH in the scope of negation weakens the meaning of the overall sentence and is thus blocked. Consider now what would happen if we were to lift the ban in (11): we could then parse the negative (2a) (repeated below), as in (39). The latter is associated with the weaker NEGATED FREE CHOICE meaning, $\neg \Diamond C \vee \neg \Diamond B$, which is true in the given target context (Angie doesn't have free choice, as she can only buy the boat).

(2a) Angie is not allowed to buy the boat or the car.

(39) not[EXH[Angie is allowed to buy the boat or the car]]
= not[Angie can choose between the two] NEGATED FC

If (39) were a viable option, then (2a) could be associated with a true reading in the context (in addition to that in (40) without EXH, which would give rise to a false meaning instead, as seen above).

(40) not[Angie is allowed to buy the boat or the car] DOUBLE PROHIBITION

On the basis of this, one might hypothesize that the reason people gave intermediate responses to the negative FC targets was that the target sentences were associated with two different readings, one true and one false, resulting in an intermediate status (see Bill et al. 2018 and Bar-Lev 2021 for a similar idea).

The problem with this line of explanation is that it readily extends to the negative disjunction case in (21a), repeated below. This is because the latter could also be analyzed as in (41), giving rise to a weaker meaning, $\neg B \vee \neg C$, which is true in the given context.

(21a) Angie didn't buy the boat or the car.

(41) not[EXH[Angie bought the boat or the car]]
not[Angie bought the boat or the car but not both] NEGATED EXCL

Therefore, simply abandoning the principle in (11) would not allow us to account for the difference between the free choice and disjunction conditions (not to mention the fact that it would leave us without an explanation for the distribution of implicatures). What's needed is to replace (11) with a principle that still disallows EXH under negation in cases like (21a), but allows it in cases like (2a). A recent proposal in the literature, Enguehard & Chemla (2021), has independently argued for a principle that achieves exactly this. Enguehard & Chemla (2021) argue that the standard formulation in (11) based on logical strength should be replaced

by a constraint based on a notion of 'connectedness', a logical notion related to monotonicity. Roughly, the principle in (42) makes a parse strongly dispreferred if it doesn't give rise to a connected meaning.

(42) Among the parses of a sentence (with or without EXH), those that result in non-connected meanings are dispreferred/marked.

Formal details aside (for which we refer the reader to Enguehard & Chemla 2021), what is relevant here is that this proposal differs from the standard one precisely in that it predicts EXH not to be banned in the scope of negation with FREE CHOICE, while still predicting the standard asymmetry between positive and negative in the case of simple disjunction.[13] Their argument is that it would intuitively be easier to force a NEGATED FREE CHOICE reading as in (43), than the corresponding negated exclusivity meaning in (44).[14]

(43) Angie is not allowed to buy the boat or the car. She's only allowed to buy the boat!

(44) Angie didn't buy the boat or the car. She bought both!

The intuitive difference between (43) and (44) is in line with what we observed in our experiment, in particular with respect to the comparison between the negative free choice and disjunction conditions. Under this hypothesis, participants would have found it easier to read the negative FC targets with an embedded EXH than to do the same for the corresponding negative OR targets. As a result, the former had a true reading in the context, and consequently led participants to choose the medium strawberry. This proposal presents a promising direction for making the implicature approach compatible with our results (perhaps in combination with considerations of scalar diversity, to account for the difference in the positive conditions).

13 This is the case only if an anti-conjunctive inference is not derived (see Enguehard & Chemla 2021).

14 Both are actually possible with marked intonation, but the claim is that (43) is possible and relatively easy without such an intonation. As Enguehard & Chemla (2021) discuss, their proposal doesn't account for why the negated free choice reading in (43), while possibly easier to access than the corresponding negated exclusivity reading, is intuitively still more difficult than the DOUBLE PROHIBITION reading in (i).

(i) Angie is not allowed to buy the boat or the car.
⇝ *Angie is not allowed to buy either one* DOUBLE PROHIBITION

9.3.2 Scope and ambiguity

Another option for the implicature approach is to reconsider the possible parse of the sentences we mentioned above in passing in the discussion of Experiment 1, in which disjunction takes wide scope.[15] That is, both (45) and (46) could in principle be interpreted with disjunction scoping over the modal (and negation), giving rise to readings which we could paraphrase as *Angie is allowed to buy the boat or she is allowed to buy the car* and *Angie is not allowed to buy the boat or she is not allowed to buy the car*, respectively.

(45) Angie is allowed to buy the boat or the car.

(46) Angie is not allowed to buy the boat or the car.

The wide scope readings of (45) and (46) would be made true in the positive and the negative FC target conditions, respectively, as the contexts were such that Angie was allowed to buy the boat but not the car. If participants had accessed these wide scope readings, they would have selected the maximal reward in these conditions, which is not what we observed. On the other hand, under the wide scope parse, (45) and (46) would also give rise to so-called 'ignorance inferences' according to which the speaker is not certain as to whether Angie is allowed to buy the car and whether she is allowed to buy the boat (see Sauerland 2004, Fox 2007, among others). One could hypothesize that these ignorance inferences were the source of the intermediate responses we observed.[16]

As discussed, however, our experimental design adopted the prediction mode, which is specifically conceived to control for ignorance inferences: the puppet was making predictions about what would end up happening, which is consistent with the puppet being ignorant as to which of the disjuncts was true. This strategy has been employed successfully before to control for ignorance in other studies (Tieu et al. 2017a, among others). In this paradigm, then, if participants interpreted the sentences with the parses suggested above, they would nevertheless be expected to select a maximal response and not an intermediate one — whether or not they also drew ignorance inferences from the sentences. For this reason, we find the story sketched above about wide scope readings and ignorance inferences being the source of the observed intermediate responses not very plausible.

There remains, nonetheless, a possible route for the wide scope explanation,

15 We thank an anonymous reviewer for pushing us to consider this possibility in detail.

16 The theoretical situation here is complicated by the fact that, under certain conditions, the wide scope interpretation of these sentences would in itself give rise to free choice, in which case the prediction for this parse would be the same as that for the narrow scope disjunction (see Goldstein 2019 and Aloni 2022, among many others). We set aside this option here, given that this is not a prediction of the implicature approach.

which we briefly mentioned in the discussion of the disjunction controls in Experiment 1. The idea there was that when participants are faced with a sentence that is ambiguous between two parses, one of which is true and the other of which is false, they might judge the sentence as having an intermediate value. One could consider a similar explanation to apply to our free choice conditions. The sentences in the target conditions are predicted to be either plainly false or associated with an implicature violation, under the narrow scope reading. However, if interpreted with the inverse scope parses suggested above, they should be associated with a maximal reward. It is therefore possible that at least some participants perceived an ambiguity between the surface and inverse scope readings, and thus assigned the sentences an intermediate reward, in line with our results.

Having said this, recall that Experiment 2b and Experiment 3, which included additional training on conjunctive sentences, yielded fewer intermediate responses to the free choice targets (and more minimal rewards). We think that this finding, incidentally, speaks against the hypothesis about scope sketched above, in that it is not clear why additional training on conjunction should lead people to change their strategy regarding scopally ambiguous sentences and the intermediate reward.

Finally, we also do not think this is how participants would judge genuinely scopally ambiguous sentences in a ternary judgment task. For instance, when presented with (47) in a context in which two out of three horses jumped over the fence, we think participants would tend to access an inverse scope reading and give the sentence a maximal reward.

(47) Every horse didn't jump over the fence.

Some suggestive evidence in this direction comes from one of the control conditions reported in Tieu et al. (2019), who used the ternary judgment task to investigate homogeneity in plural definite descriptions in French. In their Experiment 2, they included so-called scope ambiguity controls of the form *Tous les coeurs ne sont pas rouges* 'All the hearts are not red' (intended interpretation: *Not all the hearts are red*), presented in contexts in which two of four hearts were red. The adult controls in the study tended to interpret the universal as scoping under the negation, and predominantly selected the maximal reward on these trials.

10 Conclusion

A sentence containing disjunction in the scope of a possibility modal, such as *Angie is allowed to buy the boat or the car*, gives rise to the FREE CHOICE inference that Angie can freely choose between the two. As discussed, this inference is puzzling for standard treatments of modals and disjunction. In addition, FREE CHOICE tends to disappear under negation: *Angie is not allowed to buy the boat or*

the car doesn't merely convey the negation of free choice, but rather the stronger DOUBLE PROHIBITION reading that Angie cannot buy either one. There are two main approaches to capturing this pattern in the literature, one of which appeals to an implicature mechanism. While both the implicature and non-implicature approaches cover the basic pattern, and more complicated related data points, they diverge in what they predict for the status of positive and negative sentences in certain contexts. In this paper, we presented a set of four experiments testing these predictions.

Experiment 1 focused on the comparison between positive and negative free choice statements and positive and negative plain disjunctive statements. If the free choice inference is an implicature, we should expect it to behave like the exclusivity implicature of plain disjunction with respect to polarity. Contrary to this expectation, we observed a significant interaction between inference type (free choice vs. plain disjunction) and polarity (positive vs. negative): people distinguished the two polarities more for plain disjunction than they did for free choice. In Experiment 2a, we compared positive and negative free choice to the scalar inference of the modal *is allowed to*/*is not allowed to*. Here too, we observed a significant interaction between inference type (free choice vs. modal) and polarity (positive vs. negative), with people distinguishing between the polarities for the modal but not for free choice. The findings of both Experiments 1 and 2a run counter to the predictions of the scalar implicature account.

However, while Experiments 1 and 2a provided some suggestive evidence against the implicature account, we identified a potential confound: participants might have employed a charitable response strategy, offering the intermediate reward when the puppet was right about at least one of the mentioned objects. We reasoned that this strategy might be encouraged precisely by the utterance of disjunction: the puppet specifically mentioned two objects and turned out to be 'right' about one of them, hence the intermediate reward is appropriate. To address this concern, we conducted Experiment 2b, which included training items that would highlight to participants that they should not select the intermediate reward for literally false sentences, and in Experiment 3 we turned to an instance of free choice that does not involve explicitly pronouncing the individual disjuncts: free choice 'any', and we compared this to the scalar implicature of 'some'. And indeed, we saw less evidence of a charitable 'partial truth' strategy. Moreover, our false conjunctive controls allowed us to exclude participants who persisted in the charitable response strategy, despite the additional training. With these controls in place, we nevertheless again observed a significant interaction between inference type (free choice disjunction vs. the modal, free choice 'any' vs. 'some') and polarity (positive vs. negative) — counter to the predictions of the implicature account.

The results of the full set of experiments present a challenge for the implicature

approach but are more straightforwardly in line with the homogeneity account. At a more general level, our results are consistent with previous findings in the literature of crucial differences between free choice and implicatures, e.g., in their processing and acquisition profiles (Chemla & Bott 2014, Tieu et al. 2016).

One could take our results as providing support for a non-implicature approach; alternatively, the findings might encourage us to refine the implicature approach, by reconsidering, for instance, the principle regulating the distribution of the EXH operator. Regardless of the theoretical choice pursued, empirically investigating the relative status of the positive and negative sentences provides a useful way to address the debate between implicature and non-implicature approaches to free choice.

Ethics and consent

This study was approved by the University of Toronto Social Sciences, Humanities and Education Research Ethics Board (protocol no. 43403).

Funding

This work has been supported by the Social Sciences and Humanities Research Council of Canada (Insight Development Grant awarded to L. Tieu), the Connaught Fund's New Researcher Award (awarded to L. Tieu), the Leverhulme trust grant RPG-2018-425 (awarded to J. Romoli), and Western Sydney University's Research Theme Champion support funding (awarded to L. Tieu).

Data accessibility

The design files, anonymized data, and R analysis scripts can be found at the following OSF entry: https://osf.io/56qvw/?view_only=74fee94d012642609eba8663bb78d4d3.

Acknowledgements

For helpful discussion and feedback, we would like to thank Amir Anvari, Moysh Bar-Lev, Richard Breheny, Lucas Champollion, Emmanuel Chemla, Gennaro Chierchia, Milica Denic, Emile Enguehard, Danny Fox, Simon Goldstein, Martin Hackl, Nathan Klinedinst, Mora Maldonado, Matt Mandelkern, Paul Marty, Salvador Mascarenas, Paolo Santorio, Florian Schwarz, Benjamin Spector, Yasu Sudo, Alexis Wellwood, and the audiences at XPRAG 2019, SALT 29, ELM 2, and at MIT, NYU, UCSD, University of Göttingen, UCL, and University of Amsterdam.

Competing interests

The authors declare no competing interests.

References

Abrusan, Marta & Kriszta Szendroi. 2013. Experimenting with the King of France. *Semantics & Pragmatics* 6(10). 1–43.

Aloni, Maria. 2007. Free choice, modals, and imperatives. *Natural Language Semantics* 15(1). 65–94. http://dx.doi.org/10.1007/s11050-007-9010-2.

Aloni, Maria. 2022. Logic and conversation: the case of free choice. *Semantics & Pragmatics* 15. http://dx.doi.org/doi.org/10.3765/sp.15.5.

Augurzky, Petra, Marion Bonnet, Richard Breheny, Alexandre Cremers, Cornelia Ebert, Clemens Mayr, Jacopo Romoli, Markus Steibach & Yasutada Sudo. 2023. Putting plural definites into context. In *Proceedings of Sinn und Bedeutung 27*, https://ling.auf.net/lingbuzz/007106?_s=aqzVP9HC_hzOFedF&_k=UkMJGTMoeuHZao_3.

Bar-Lev, Moshe. 2018. *Free choice, homogeneity and innocent inclusion*: The Hebrew University of Jerusalem dissertation.

Bar-Lev, Moshe. 2021. An implicature account of homogeneity and non-maximality. *Linguistics & Philosophy* 44(5). 1045–1097. http://dx.doi.org/10.1007/s11050-015-9114-z.

Bar-Lev, Moshe & Danny Fox. 2017. Universal free choice and innocent inclusion. In *Proceedings of SALT 27*, 95–115.

Bar-Lev, Moshe & Danny Fox. 2020. Free choice, simplification, and Innocent Inclusion. *Natural Language Semantics* 28. 175–223. http://dx.doi.org/10.1007/s11050-020-09162-y.

Barker, Chris. 2010. Free choice permission as resource-sensitive reasoning. *Semantics and Pragmatics* 3(10). 1–38.

Beaver, David. 2001. *Presupposition and Assertion in Dynamic Semantics*. Stanford University: CSLI Publications.

Bill, Cory, Jacopo Romoli & Florian Schwarz. 2018. Processing presuppositions and implicatures: Similarities and differences. *Frontiers in Communication* 3(44).

Chemla, Emmanuel. 2009. Universal implicatures and free choice effects: Experimental data. *Semantics and Pragmatics* 2(2).

Chemla, Emmanuel. 2010. Similarity: towards a unified account of scalar implicatures, free choice permission and presupposition projection. Ms., Ecole Normale Supérieure.

Chemla, Emmanuel & Lewis Bott. 2014. Processing inferences at the semantics/pragmatics frontier: disjunctions and free choice. *Cognition* 130(3). 380–396.

Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention*. Oxford University Press.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning volume 3*, Berlin: Mouton de Gruyter.

Del Pinal, Guillermo, Itai Bassi & Uli Sauerland. 2022. Free choice and presuppositional exhaustification. Unpublished manuscript UMass and ZAS.

Enguehard, Emile & Emmanuel Chemla. 2021. Connectedness as a constraint on exhaustification. *Linguistics & Philosophy* 44. 79–112.

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and Implicature in Compositional Semantics*, 71–120. Palgrave.

Fox, Danny. 2012. Presupposition projection from quantificational sentences: trivalence, local accommodation, and presupposition strengthening. In Ivano Caponigro & Carlo Cecchetto (eds.), *From grammar to meaning: the spontenous logicality of language*, 201–232. Cambridge University Press.

Fox, Danny & Benjamin Spector. 2018. Economy and embedded exhaustification. *Natural Language Semantics* 26(1). 1–50.

Franke, Michael. 2011. Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics* 4(1). 1–82. http://dx.doi.org/10.3765/sp.4.1.

Fusco, Melissa. 2015. Deontic modality and the semantics of choice. *Philosophers' Imprint* 15.

Goldstein, Simon. 2019. Free choice and homogeneity. *Semantics and Pragmatics* 12(23). 1–47. http://dx.doi.org/10.3765/sp.

Gotzner, Nicole, Jacopo Romoli & Paolo Santorio. 2017. More free choice and more inclusion: An experimental investigation of free choice in non-monotonic environments. Ms. Ulster University, ZAS, and UCSD.

Gotzner, Nicole, Jacopo Romoli & Paolo Santorio. to appear. Choice and prohibition in non-monotonic contexsts. *Natural Language Semantics* .

Huang, Haiquan & Stephen Crain. 2019. When OR is assigned a conjunctive inference in child language. *Language acquisition* 74–97.

Kamp, Hans. 1974. Free choice permission. *Proceedings of the Aristotelian Society* 74. 57–74.

Katsos, Napoleon & Dorothy VM Bishop. 2011. Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition* 120(1). 67–81.

Klinedinst, Nathan. 2007. *Plurality and possibility*: UCLA dissertation.

Kratzer, Angelika & Junko Shimoyama. 2002. Indeterminate pronouns: The view

from Japanese. In Yukio Otsu (ed.), *Proceedings of the Tokyo conference on psycholinguistics*, vol. 3, 1–25. Tokyo: Hituzi Syobo.

Križ, Manuel. 2015. *Aspects of homogeneity in the semantics of natural language*: University of Vienna dissertation.

Križ, Manuel & Emmanuel Chemla. 2015. Two methods to find truth-value gaps and their application to the projection problem of homogeneity. *Natural Language Semantics* 23. 205–248.

Marty, Paul & Jacopo Romoli. 2019. Presupposed free choice and the theory of scalar implicatures. Ms., ZAS and Ulster University.

Marty, Paul, Jacopo Romoli & Paolo Santorio. 2020. Counterfactuals and Undefinedness: homogeneity vs supervaluations. In s (ed.), *Proceedings of SALT30*, 208–227. https://semanticsarchive.net/Archive/jgwZDc2Z/conditionals.pdf.

Marty, Paul, Jacopo Romoli, Yasutada Sudo & Richard Breheny. 2021. Negative free choice. *Semantics and Pragmatics* 14. 13.

Meyer, Marie-Christine. 2018. An apple or a pear: free choice disjunction. In H. Rullmann T.E. Zimmermann L. Matthewson, C. Meier (ed.), *Wiley's semantics companion*, Wiley and Sons.

Renans, Agata, Jacopo Romoli, Maria-Margarita Makri, Lyn Tieu, Hanna de Vries, Raffaella Folli & George Tsoulas. 2018. The abundance inference of pluralised mass nouns is an implicature: Evidence from Greek. *Glossa* 3(1). 103.

Romoli, Jacopo & Paolo Santorio. 2019. Filtering free choice. *Semantics & Pragmatics* (in press).

Rothschild, Daniel & Stephen Yablo. 2018. Permissive updates. Ms., UCL and MIT.

Santorio, Paolo & Jacopo Romoli. 2017. Probability and implicatures: A unified account of the scalar effects of disjunction under modals. *Semantics & Pragmatics* 10(13).

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27(3). 367–391.

Skordos, Dimitrios, Roman Feiman, Alan Bale & David Barner. 2020. Do children interpret "or" conjunctively? *Journal of Semantics* 37(2). 247–267.

Starr, Will. 2016. Expressing permission. In *Proceedings of SALT 26*, 325–349.

Sun, Chao, Richard Breheny & Daniel Rothschild. 2019. Exploring the existential/universal ambiguity in singular donkey sentences. In *Proceedings of Sinn und Bedeutung 24*, .

van Tiel, Bob, Emiel van Miltenburg, Natalia Zevakhina & Bart Geurts. 2016. Scalar diversity. *Journal of Semantics* 33(1). 137–175.

Tieu, Lyn, Cory Bill, Jacopo Romoli & Stephen Crain. 2017a. Plurality inferences are scalar implicatures: evidence from acquisition. Ms., Maquarie University and Ulster University.

Tieu, Lyn, Manuel Kriz & Emmanuel Chemla. 2019. Children's acquisition of

homogeneity in plural definite descriptions. *Frontiers in Psychology* .

Tieu, Lyn, Jacopo Romoli, Peng Zhou & Stephen Crain. 2016. Children's knowledge of free choice inferences and scalar implicatures. *Journal of Semantics* 33(2). 269–298.

Tieu, Lyn, Kazuko Yatsushiro, Alexandre Cremers, Jacopo Romoli, Uli Sauerland & Emmanuel Chemla. 2017b. On the role of alternatives in the acquisition of simple and complex disjunctions in french and japanese. *Journal of Semantics* 34(1). 127–152. http://dx.doi.org/10.1093/jos/ffw010. +http://dx.doi.org/10.1093/jos/ffw010.

Van Tiel, Bob. 2012. Universal free choice? In Ana Aguilar et al. Guevara (ed.), *Proceedings of sinn und bedeutung 16*, 627–638.

Willer, Malte. 2017. Widening free choice. In *Proceedings of the 21st amsterdam colloquium*, 511–520.

Zehr, Jérémy. 2014. *Vagueness, presupposition and truth-value judgments*: Institut Jean Nicod, Ecole Normale Supérieure dissertation.

Zimmerman, Thomas Ede. 2000. Free choice disjunction and epistemic possibility. *Natural Language Semantics* 8(255–290).