Linguistics and Philosophy
https://doi.org/10.1007/s10988-023-09381-6

ORIGINAL RESEARCH

Scalar implicatures with discourse referents: a case study on plurality inferences

Yasutada Sudo[1]

Accepted: 3 January 2023

Abstract

This paper explores the idea that scalar implicatures are computed with respect to *discourse referents*. Given the general consensus that a proper account of pronominal anaphora in natural language requires discourse referents separately from the truth-conditional meaning, it is naturally expected that the anaphoric information that discourse referents carry play a role in the computation of scalar implicatures, but the literature has so far mostly exclusively focused on the truth-conditional dimension of meaning. This paper offers a formal theory of scalar implicatures with discourse referents couched in dynamic semantics, and demonstrates its usefulness through a case study on the plurality inferences of plural nouns in English.

Keywords Discourse referents · Scalar implicatures · Dynamic semantics · Plurality inferences

The first version of the work reported here was developed for the fifth meeting of London Semantics Day held at Queen Mary University of London on 15 May 2018, and I was very fortunate to have opportunities to present various stages of it at many venues: Universiteit Utrecht (11 June 2018), Leibniz-Zentrum Allgemeine Sprachwissenschaft (15 June 2018), Ecole Normale Supérieure (21 March 2019), University of Vienna (20 May 2019), Heinrich-Heine-Universität Düsseldorf (26 June 2019), University of Oxford (5 June 2019), Goethe-Universität Frankfurt (11 July 2019), Georg-August-Universität Göttingen (6 September 2019), University of Toronto (2 December 2019), and Massachusetts Institute of Technology (6 December 2019). The feedback and encouragements I received from the audiences of these events had tremendous contribution to the quality of the present work, and I would like to thank them. I also thank the participants of the semantics seminar I co-taught with Gennaro Chierchia, Nathan Klinedinst, and Daniel Rothschild at University College London in Autumn 2019, where I discussed the material reported here. My special thanks go to Brian Buccola, Gennaro Chierchia, Kai von Fintel, Martin Hackl, Natalia Ivlieva, Filipe Kobayashi, Manuel Križ, Jeremy Kuhn, Matt Mandelkern, Rick Nouwen, Jacopo Romoli, Benjamin Spector, Ede Zimmermann, and an anonymous reviewer for *Linguistics and Philosophy*. Part of this work was funded by the AHRC-DFG Research Grant "Interactions between Dynamic Effects and Alternative-Based Inferences in the Study of Meaning" (AH/V003526/1). I am solely responsible for the remaining errors.

✉ Yasutada Sudo
y.sudo@ucl.ac.uk

[1] University College London, 2 Wakefield St., London W1CN 1PF, UK

Published online: 22 May 2023

1 Introduction

It is generally agreed that pronominal anaphora cannot be explicated solely in terms of truth-conditions (Karttunen, 1976; Kamp, 1981; Heim, 1982). This is most acutely evidenced by pairs of sentences that have contextually equivalent truth-conditions but nonetheless differ with respect to discourse anaphora. Concretely, consider the examples in (1). The pronoun *it* can refer to Paul's Taxpayer Identification Number (TIN) in (1a), but not in (1b), even if it is commonly known that every registered taxpayer has a unique TIN, and everyone who has a TIN is a registered taxpayer.

(1) a. Paul has a TIN. But he hasn't used it in a while.
b. Paul is a registered taxpayer. ??But he hasn't used it in a while.

In order to explain this observation, theories of discourse anaphora commonly postulate *discourse referents*.[1] Discourse referents are abstract semantic objects that can be introduced to linguistic discourse in certain specific ways, e.g. by using an indefinite noun phrase.[2] They represent information about what kind of entities are being talked about, and by assumption, are essential in resolving pronominal anaphora. Putting aside formal details for now, this idea explains the above contrast roughly as follows. The first sentence of (1a) introduces a discourse referent representing Paul's TIN and the pronoun can be successfully resolved to it, while the first sentence of (1b) does not introduce a discourse referent, so the pronoun in the second sentence cannot be interpreted as referring to Paul's TIN.

Although a number of different formal implementations of discourse referents have been put forward, I believe it has been uncontroversial, ever since the need for discourse referents was originally recognised more than half a century ago (see the citations at the beginning), that a proper characterisation of the meaning of natural language sentences needs to postulate (at least) two separate dimensions of meaning: discourse referents and truth-conditions.[3] Given this background, it is quite surprising to notice that potential implications of the multi-dimensionality of natural language semantics on pragmatic inferences have so far not been given enough attention in the theoretical

[1] The term *discourse referent* tends to be associated with (certain versions of) dynamic semantics, but the idea of discourse referents as carriers of anaphoric potentials is in fact more general. This is explained in more detail in Appendix A.

[2] Indefinites are often used as examples of means to introduce discourse referents, but as Heim (1982) emphasises, there are other ways of introducing discourse referents too, including non-linguistic means. Later in this paper I will adopt the view put forward by Van den Berg (1996) among others that all quantifiers, including indefinites, introduce discourse referents.

[3] Natural language semantics is believed to be multidimensional in other respects as well. Most notably, so-called *expressive meaning* is often considered to be semantically independent from truth-conditions (Potts, 2005, 2007), and perhaps also from the anaphoric dimension of meaning. But some recent work on expressives (McCready, 2010; Gutzmann, 2015; Gutzmann and McCready, 2016; among others) propose to blur the boundary between the expressive and truth-conditional dimensions. Presupposition is more controversial: while many regard presuppositions to be part of truth-conditions (Heim, 1982, 1983; Beaver, 2001; Beaver and Krahmer, 2001; George, 2008; Fox, 2012), some argue that they should be treated as a separate dimension of meaning (Karttunen and Peters, 1979; Sudo, 2012, 2014). If either or both of these phenomena call for a different dimension of meaning, their behaviour with respect to pragmatic inferences should also be examined in detail, but as far as I know, this question has not been investigated in sufficient detail, and is left unaddressed here as well.

literature. Among various pragmatic inferences, I will focus on *scalar implicature* in this paper, and discuss what roles discourse referents can and should play in this phenomenon.

Scalar implicature has been very actively studied since the 1970s, and consequently the literature is extremely copious, but discourse referents have been almost always ignored (exceptions include Geurts, 2008, 2009; Sudo, 2016).[4] I would say this negligence in the literature is unjustifiable. Virtually all theories of scalar implicature, although technically and conceptually diverse, share the fundamental insight that traces back to Grice (1989), which roughly goes as follows: If sentence ϕ has a (contextually relevant) alternative ψ that is more informative (or is not less informative, according to some theories), then an utterance of ϕ will have a scalar implicature that amounts to the negation of ψ. In most current implementations of this idea, informativity is understood solely in terms of truth-conditions. That is, sentence ψ is said to be more informative than sentence ϕ, if whenever ψ is true, ϕ is true, but not vice versa. However, the idea of informativity itself is more general than just this, and applicable to any type of information. As mentioned above, it is agreed that discourse referents represent information distinct from the truth-conditional aspect of meaning, then it makes sense to also speak of the informativity of discourse referents introduced in ϕ and ψ. That is, if it so happens that an alternative ψ of some sentence ϕ introduces a more informative discourse referent than ϕ does, then it is expected that an utterance of ϕ will give rise to some scalar implicature that amounts to the 'negation' of this extra bit of information that the discourse referent of ψ would carry, if ψ had been uttered instead. This is exactly the idea I would like to explore in this paper.

I will take one empirical phenomenon, namely, the plurality inferences of plural noun phrases in English, as a case study, and claim that understanding plurality inferences as scalar implicatures that involve discourse referents allows us to achieve a straightforward analysis of this phenomenon. I will offer one particular way of formalising the analysis in a (relatively plain) version of *dynamic semantics*. My choice of framework is not theoretically very crucial, but practically motivated: Dynamic semantics is arguably the most thoroughly worked out formal theory of discourse referents as of today. Also, the fact that interactions between quantification and discourse referents have been extensively discussed by previous authors is a big advantage of this framework for my purposes here. I will remark on these points in more concrete terms as we go along.

Having set the goal, I should also mention that it is not my aim in this paper to argue against other theories of plurality inferences. In particular, the empirical predictions of my theory will be very close, though not identical to, to those of Spector (2007). However, as I will argue below, my analysis will be conceptually more parsimonious in that it will allow us to dispense with certain crucial extra assumptions Spector (2007)

[4] One of the prominent views of scalar implicature in the current theoretical literature is the so-called *grammatical approach*, according to which scalar implicatures are not pragmatic inferences in the classical sense—in the sense of involving post-semantic reasoning about conversational intentions—but are actually generated by a phonologically silent linguistic operator [see Chierchia et al. (2012) and references therein]. However, it is still true that discourse referents have been largely neglected in this corner of the literature as well. A version of the grammatical approach to scalar implicature that makes reference to discourse referents will be presented in Sect. 4.3.

and others of the same persuasion make about scalar alternatives to plural nouns. In addition, towards the end of the paper, I will demonstrate that my theory leads to a new way of understanding an empirical observation made by Crnič et al. (2015) about the so-called distributivity inference of disjunction under a universal quantifier (see also Bar-Lev and Fox, 2020). Moreover, I could certainly eventually be shown to be on the wrong track for this particular linguistic phenomenon, but I believe the formal theory that I develop here has independent theoretical value as a proof of concept for the idea of scalar implicatures with discourse referents, since, as I have already remarked, as long as we follow Grice's insights, discourse referents must be relevant for scalar implicature, and this is the first systematic study that explores this idea.

The present paper is structured as follows. In Sect. 2, I will review the main empirical phenomenon, the plurality inferences of plural nouns phrases in English, as well as the theoretical landscape in the current literature on this phenomenon, and sketch my proposal in informal terms. After introducing a simple dynamic semantic system in Sect. 3, I will show in Sect. 4 a concrete formal implementation of my theory of plurality inferences, including the details of how scalar implicatures are to be computed with respect to discourse referents. Then in Sect. 5, I will extend the analysis to cases involving quantifiers by incorporating so-called externally dynamic selective generalised quantifiers in the system, and discuss a consequence of this extension on the distributivity inference of disjunction under a universal quantifier in Sect. 6. Finally, I will conclude in Sect. 7.

2 Plurality inferences

Terminologies in linguistics can be quite misleading. The term *plural* for certain forms of nouns in languages like English is one example of this. To illustrate, consider the following sentences.

(2) a. This coat has pockets.
    b. This coat does not have pockets.

A noun like *pockets* is standardly called plural, and this is presumably because a sentence like (2a) has a very robust *plurality inference* that the coat in question has more than one pocket. However, what is surprising is that the negation of this sentence, (2b), does not mean the negation of (2a) with the plurality inference, which would be 'The coat does not have multiple pockets'. Rather, the observed meaning of (2b) is stronger than this, namely, that the coat has no pocket whatsoever.

This is part of the main empirical puzzle we will be concerned with in this paper. More generally, plural nouns like *pockets* in certain sentences like the negative sentence in (2b) do not behave as would be expected if they had plural meaning and obeyed the principles of compositional semantics, and this is why it is misleading to call nouns like *pockets* 'plural'.[5]

[5] As Spector (2007) points out, plural nouns under negation are actually not completely number-neutral, because otherwise the infelicity of (i) would be unexpected.

A natural question to ask in light of the above observation is when a plurality inference is observed and when it is not. As previous studies have uncovered, the overall generalisation is that bare plurals like *pockets* give rise to number-neutral readings in negative contexts that largely overlap with implicature cancelling contexts (but see Grimm, 2013 for potential issues). These include sentences with sentential negation like (2b) above, as well as those in (3).[6]

(3) a. If you have *coins* in your pocket, please put them in the tray.
b. We should clean up the mess before *guests* arrive.
c. This plant can survive without *leaves* for many years.

A number of different analyses have been proposed to account for the distribution of plurality inferences. Here is a short summary of the current literature.

- The scalar implicature approach (Mayr, 2015; Spector, 2007; Ivlieva, 2013, 2020; Zweig, 2009) analyses plurality inferences as scalar implicatures.
- The anti-presupposition approach (Sauerland, 2003; Sauerland et al., 2005) derives plurality inferences as 'anti-presuppositions'.
- The ambiguity approach (Farkas and de Swart, 2010; Grimm, 2013; Martí, 2020) postulates ambiguity between plural and number-neutral meaning.
- The homogeneity approach (Križ, 2017) likens the semantic behaviour of bare plurals to that of definite plurals, which are known to exhibit homogeneity effects.

Of these, the anti-presupposition approach is historically the oldest, but at least the versions proposed by Sauerland (2003) and Sauerland et al. (2005) are known to have a serious empirical flaw with respect to quantification. As Spector (2007) discusses

Footnote 5 continued

(i) #My son doesn't have Roman noses.

I do not think there is a very insightful account of this observation in the current literature. Spector (2007) mentions it essentially as an open problem for his account. Farkas and de Swart (2010) discuss it in some detail but do not give a satisfactory explanation. It seems to me that this is actually a probabilistic phenomenon in the sense that the anomaly of (i) comes from the inference that the probability (according to the common ground) of my son's having multiple noses is high. The reason why I think this inference is probabilistic is because of cases that do not have extreme probabilities. For example, consider (ii).

(ii) a. The postdoc will not submit an abstract to CUNY.
b. The postdoc will not submit abstracts to CUNY.

Both of these are acceptable sentences, but they are not completely equivalent. That is, if the probability of the postdoc submitting exactly one abstract is deemed high, (iia) is the more natural choice, but if it's more likely that they will submit multiple abstracts, (iib) sounds better. Also, these judgments become sharper when the relevant probability approaches either extreme. This aspect of meaning is highly intriguing but at this point it requires more careful empirical research. Since it is not immediately relevant for this paper, I will leave it for future research.

[6] Bare plurals in polar questions like *Does this coat have pockets?* are often mentioned in this connection, but their judgments seem to be more involved than often assumed (see Pearson et al., (2010); Bale et al. (2011) for related remarks). In particular, whether the *yes*-answer is felicitous when the coat only has one pocket seems to be affected by extra-linguistic considerations as well, e.g. whether or not the speaker needs more than one pocket to achieve what they want to achieve. This could be due to the semantics/pragmatics of the answer particles, or the semantics/pragmatics of polar questions, or both. Either way, this will not directly matter for my theory of plurality inferences.

this problem in detail, I will not delve into it here. Among the other three, I will adopt the scalar implicature approach in this paper, because I believe its empirical coverage is broader than the other two, especially with respect to what I call *partial plurality inferences*.

2.1 Partial plurality inferences

There are two types of partial plurality inference. The first kind, exemplified by (4), is discussed by Spector (2007) (see also Ivlieva, 2014; Križ, 2017).[7]

(4) Exactly one of these coats has pockets.

This sentence has a plurality inference in the sense that it implies that the unique coat that has pockets has multiple pockets. However, this inference is only partially plural because with respect to the other coats, *pockets* is understood number-neutrally, as the sentence entails that these coats do not have any pocket, rather than merely that they do not have multiple pockets. Note that the fully plural reading might also be available for this sentence, but what is of interest here is the reading that is stronger on the negative side of the meaning.

The second kind of partial plurality inference involves a definite plural with a bound pronoun and is discussed by Sauerland (2003) and Sauerland et al. (2005). Consider (5), for example.

(5) Every passenger of this flight lost their suitcases.

This sentence has a presupposition that makes it infelicitous if every passenger has exactly one suitcase. Crucially, this presupposition does not require every passenger to have multiple suitcases, but rather only that at least some of the passengers have

[7] In principle the same point could be made with a plural quantifier like *exactly two*, *an odd number of*, etc., as in (i), but it would allow for a dependent plural reading of the plural object, which would muddle the judgments.

(i) Exactly two of my students wrote interesting papers.

As Ivlieva (2013, 2020) discusses in great detail, dependent plurals give rise to very interesting issues. Above all, as she points out, they cannot be reduced simply to cumulative readings (*pace* Zweig, 2009). Furthermore, Ivlieva (2014, fn. 2) points out that Spector's (2007) version of the scalar implicature approach fails to derive them straightforwardly. I believe my account of plurality inferences developed in this paper has some bearing on how to explain dependent plurals, but I would like to address it in a separate occasion, as we would need non-distributive quantifiers to give a complete account of dependent plurals, which would make this paper at least twice as long.

Another type of quantifier I could use here to illustrate partial plurality inferences is a universal quantifier like *every*.

(ii) Every coat has pockets.

As discussed by Spector (2007), Zweig (2009) and Križ (2017), this sentence seems to have a partial plurality inference that at least some of the coats have multiple pockets. However, arguably, it also has a prominent reading that entails that every coat has multiple pockets, so the judgments are perhaps less clear (cf. Ivlieva, 2020: fn. 32).

multiple suitcases. Thus, this plurality inference is in the presuppositional domain and is partial in the sense that it does not apply to every passenger.

These partial plurality inferences pose issues for certain approaches to plurality inferences. Farkas and de Swart (2010), who put forward an ambiguity theory, explicitly acknowledge that sentences like (4) pose a significant challenge for their ambiguity theory. They do not mention examples like (5), but such examples are equally problematic for their account. Here is why. They postulate two meanings for each plural noun at the lexical level, a semantically plural meaning and a number-neutral meaning, and put a constraint on their distributions so as to explain why simple sentences like (2a) and (2b) are not perceived as ambiguous. The problem is that under this view, there is no way to simultaneously assign both plural and number-neutral meanings to a single occurrence of a plural noun, but that is exactly what would be needed to account for the partial plurality inferences.

I believe that partial plurality inferences are potentially problematic for Križ's (2017) homogeneity approach as well, but explaining why will require a rather long detour, so I will spell it out in Appendix B.

This leaves us with the scalar implicature approach. As we will discuss below, at least certain version of this approach can deal with both types of partial plurality inferences. However, there's a drawback, at least at the conceptual level: Existing theories that take the scalar implicature approach rely on certain additional theoretical machinery. I will not delve into all the technical bolts and nuts of these different theories, but I will point out that the idea of scalar implicatures triggered in reference to discourse referents allows us to dispense with such additional mechanisms.

2.2 The scalar implicature approach

The core assumptions of the scalar implicature approach to plurality inference are (i) that plural nouns are semantically number-neutral, and (ii) that a plurality inference arises from a plural noun as a scalar implicature via competition with its singular counterpart. (i) leads to a straightforward account of the number-neutral interpretation of plural nouns in negative contexts, while (ii) is meant to account for occurrences in positive contexts like (2a), repeated here as (6a). More specifically, the plurality inference of this example is generated in reference to its alternative in (6b), which has a singular noun in place of the plural noun.

(6) a. This coat has pockets.
    b. This coat has a pocket.

Although the idea that the plural competes with the singular is not at all inconceivable, and even shared by certain other theories (e.g., Farkas and de Swart, 2010), there is an issue here. Under the assumption that the plural noun is semantically number-neutral, the two sentences in (6) will come out as truth-conditionally equivalent. Specifically, it is clear that whenever (6a) is true, (6b) will be true. Furthermore, whenever (6b) is true, there must be at least one pocket on the coat, and this one pocket will be enough to make (6a) true, since the plural noun is assumed to be semantically number-neutral. This truth-conditional equivalence of the two sentences is an issue for the scalar implicature approach, because in order to start the computation of a scalar implicature, there needs to be some semantic asymmetry between the two sentences.

Different solutions to this issue can be found in the literature. Spector (2007) employs *higher-order implicatures*. Putting details aside, Spector's idea applied to (6a) amounts to that its crucial alternative is not (6b) on its literal reading, but on the reading that is enriched with its own scalar implicature. Note that (6b) can have a reading that implies that the coat in question has only one pocket. Since this alternative is truth-conditionally more informative than (6a), (6a) will have a scalar implicature that the alternative is false, which, together with the literal meaning of (6a), implies that the coat has multiple pockets.

Ivlieva (2013, 2014, 2020), Mayr (2015) and Zweig (2009) pursue a different solution that resorts to *embedded implicatures*. Their crucial observation is that while the two sentences in (6) are indeed truth-conditionally equivalent, the words and phrases that make them up are not, and one can find sub-constitutes of these sentences that have different truth-conditional meanings. Assuming that scalar implicatures can be drawn at the level of such sub-constituents, the plurality inference of (6a) can be computed as an embedded implicature. The authors cited here make use of embedded implicatures drawn at different constituents, but these details are not very important for the current discussion.

At this point, let me briefly show how the scalar implicature approach can deal with the two types of partial plurality inferences. It turns out that no previous analysis of the kind that uses embedded scalar implicatures offers a complete explanation of (4), as pointed out by Ivlieva (2013, 2014), so I will spell out the analysis using Spector's (2007) higher-order implicature theory here.[8]

Firstly, to derive the partial plurality inference for (4), observe first that the number-neutral semantics of the plural noun already correctly captures the negative part of the meaning. Then all we need is a scalar implicature that implies that the unique coat that has a pocket has multiple pockets. The crucial alternative is the version of the same sentence with a singular noun in place of the plural noun, as in the case of simpler sentences. Spector (2007) assumes that this alternative itself can have a scalar implicature, based on an alternative to it that contains *several pockets* in place of *a pocket*. Together with this scalar implicature, the singular alternative means (7), where the second conjunct is the negation of the meaning of the alternative "Exactly one of the coats has several pockets".

(7) Exactly one of the coats has one or more pockets, and it is false that exactly one of the coats has multiple pockets.
= Exactly one of the coats has a pocket and it has only one pocket.

Since the literal meaning of (4) is truth-conditionally equivalent to the first part of (7), (7) is truth-conditionally more informative than (4). As a result, (4) will have a scalar implicature that (7) is false. Conjoining this scalar implicature with the literal

[8] Adopting an operator-based grammatical theory of scalar implicatures, Ivlieva (2014) proposes a solution with two crucial assumptions that are yet to be given independent motivation. Specifically, her solution requires the following two stipulations: (i) when embedded, the operator that computes scalar implicatures only affects the alternatives but not the prejacent, and (ii) certain alternatives must be pruned under certain very specific circumstances.

meaning of (4), we obtain the overall meaning that implies that the unique coat that has a pocket has multiple pockets.

Similarly, the partial plurality inference of (5) can be accounted for as follows. This is a case of scalar inferences in the dimension of presuppositional meaning, and depending one's analysis of such inferences, it is perhaps not to be accounted for as a scalar implicature, but the core insight of the scalar implicature approach carries over to this case straightforwardly. Specifically, the number-neutral semantics of the plural noun *suitcases* predicts that (5) presupposes that every passenger has at least one suitcase, as presuppositions generally project universally through a universal quantifier (see, e.g. Heim, 1983; Chemla, 2009; Sudo, 2012, 2014; but see Beaver, 2001; Beaver and Krahmer, 2001; George, 2008; Fox, 2012 for different views). Now, compare this example to the version of the sentence with the singular noun *suitcase* in place of the plural noun, (8).

(8) Every passenger lost their suitcase.

This sentence presupposes that every passenger has exactly one suitcase, which comes from the uniqueness presupposition of the definite singular noun together with presupposition projection through the universal quantifier. Now we assume that a scalar inference can be drawn in the domain of presupposition as well by a mechanism similar to how scalar implicatures are computed, as suggested in the literature (Heim, 1991; Percus, 2006; Sauerland, 2008; Gajewski and Sharvit, 2012; Schlenker, 2012; Spector and Sudo, 2017; Marty, 2017; Anvari, 2019). Then, since the presupposition of (8) is stronger than the presupposition of (5) (while their at-issue meanings are equivalent), the latter comes to have the scalar inference that the presupposition of (8) is not met. This captures the fact that (5) is infelicitous when every passenger has exactly one suitcase.[9]

In sum, the scalar implicature approach is the only approach in the current literature that can deal with both types of partial plurality inference, but previous implementations of it crucially rely on an additional mechanism, namely, either higher-order implicatures or embedded implicatures. I do not think these mechanisms are unfounded or empirically problematic. In fact, embedded implicatures have been given various empirical support (see, e.g. Chierchia et al., 2012), although controversies persist in the experimental literature (Geurts and Pouscoulous, 2009; Clifton and Dube, 2010; Chemla and Spector, 2011; Geurts and van Tiel, 2013; Cummins, 2014; van Tiel, 2014; Potts et al., 2016; Franke et al., 2017; van Tiel et al., 2018). Also, the idea of higher-order implicatures seems to me to be conceptually very natural, especially if scalar implicatures are to be understood as pragmatic inferences in the Gricean sense. I will argue below, however, that as far as plurality inferences are concerned, there is no need for such additional mechanisms, once we recognise the possibility that scalar implicatures can be drawn from scalar implicatures.

[9] Note that the resulting inference is weaker than the presupposition that at least one passenger has multiple suitcases, because it only anti-presupposes the presupposition of (8), i.e. the derived inference is that the presupposition of (8) is not satisfied in the current context, rather than that it is presupposed in the current context that the presupposition of (8) is false. See Chemla (2008) for an idea that such an anti-presupposition can be pragmatically strengthened.

Before moving on, I would like to quickly remark on a potential objection against the scalar implicature approach to plurality inferences, namely that a plural inference feels much more robust than the typical scalar implicature. In particular, it does not seem to be possible to explicitly cancel a plurality inference, as illustrated by (9).

(9) This coat has pockets. ??In fact, it has only one.

In comparison, it is often considered that something likemore or less acceptable.

(10) She solved some of the problems. In fact, she solved them all.

The robustness of plurality inferences certainly needs to be explained one way or another, and I admit that I do not have much to offer here, but it is too hasty to conclude from this observation alone that plurality inferences are not scalar implicatures. In particular, recent experimental research on scalar implicatures reveals that different scalar items have scalar implicatures to different degrees of robustness (e.g., van Tiel et al., 2016, 2019; Meyer and Feiman, 2021; van Thiel and Pankratz, 2021; Marty et al., 2022; see also Singh, 2019; Bar-Lev and Fox, 2020 for theoretical discussion). I do not have anything insightful to say about this poorly understood issue of diversity across scalar items in the present paper, but it is theoretically possible that the plurality inference is a very robust type of scalar implicature.[10] I would also like to re-emphasise that the scalar implicature approach to the plurality inference currently enjoys empirical superiority over its alternatives, especially with respect to partial plurality inferences.

2.3 A rough sketch of the proposal

My version of the scalar implicature approach to plurality inferences makes crucial use of discourse referents. I will first sketch the idea with the same examples as above. In (11), discourse referents are explicitly marked.[11]

[10] Furthermore, there is some experimental data suggesting that the plurality inference is not as robust as plain entailments and is more like a scalar implicature. For instance, Anand et al. (2011) report an experiment (Experiment 4) where they tested the acceptability of sentences like (i) in contexts where there is only one entity that can be described by a plural noun.

(i) The baby wearing yellow is playing with teddy bears.

In their results, the sentences were accepted 17% of the time (as opposed to 98% when there are more than one entity describable by the plural noun), indicating that the reading without a plurality inference is not completely inaccessible, although it needs to be tested against a reasonable baseline in order for it to be more convincing evidence. Similarly, Tieu et al. (2020) tested similar sentences in their acquisition studies. The adult participants of their Experiment 2 accepted them 25% of the time, which was comparable to the acceptability of the scalar implicature triggered by *some*, although in their Experiment 1, such sentences were only accepted 8% of the time by their adult participants, although, again, these numbers need to be understood against reasonable baselines. It should also be mentioned that acquisition data point to the conclusion that plurality inferences are similar to scalar implicatures (Pearson et al., 2010; Tieu et al., 2020; Tieu and Romoli, 2019 for an overview).

[11] Here and throughout, I will use the Barwise notation where a newly introduced discourse referent is represented as a superscript, and anaphoric dependency to an old discourse referent is represented as a subscript. I will also assume without argument that discourse referents are introduced by noun phrases, but this assumption is not crucial. What is crucial here is the anaphoric properties at the sentential level, or

(11) a. This coat has [pockets]^x. (But they$_x$ are not very big.)
b. This coat has [a pocket]^x. (But it$_x$ is not very big.)

Since these sentences do feed pronominal anaphora in subsequent discourse, as indicated in parentheses, there is evidence for the discourse referents.[12] I follow the previous proponents of the scalar implicature approach and assume that a singular noun is only true of atomic entities, while a plural noun is semantically number-neutral. Recall that this assumption renders the above two sentences truth-conditionally equivalent, which, as we saw, was an issue for the scalar implicature approach, because in order to generate a scalar implicature there needs to be some semantic asymmetry between them.

My version of the scalar implicature theory distinguishes itself from its predecessors in that it finds the crucial semantic asymmetry in the discourse referents of the sentences, rather than in their truth-conditions. That is, given the above assumptions about the nominal semantics, the discourse referent introduced in (11a) is not specified for number, so x can refer to an atomic pocket or a plurality of pockets, while the discourse referent introduced in (11b) is specified to be singular, referring to an atomic pocket. Since the latter discourse referent is more informative in the sense that it carries more precise information about what it represents, a scalar implicature is drawn for the former that whatever the latter means is not the case. This ultimately amounts to restricting the possible referents of x in (11a) to non-atomic entities, which is the plurality inference.

I skipped many details above, including important questions about how exactly to define informativity for discourse referents and how to actually draw a scalar implicature from the information carried by discourse referents. I will make these points formally more precise below by implementing the idea sketched above in dynamic semantics, and then I will show how the resulting theory accounts for interactions with quantifiers, including cases of partial plurality inferences.

As we will see, the explanation for the partial plurality inference of (5) will be essentially identical to what we reviewed above for Spector's (2007) theory, but the analysis of the partial plurality inference of (4) under the present account is worth characterising in informal terms here. Specifically, the relevant two sentences are as follows. The crucial observation here is that they do introduce discourse referents about pockets, as evidenced by the continuations in parentheses.

(12) a. Exactly one of these coats has pockets^x. (And they$_x$ are big.)
b. Exactly one of these coats has [a pocket]^x. (And it$_x$ is big.)

Here again, we reason about what x can be. Given the semantic assumptions about nominal number, x in (12a) can be an atomic pocket or a plurality of pockets. On the

Footnote 11 continued
more precisely, at the level where the plurality inference is computed. Sub-sentential compositionality is obviously an important and interesting issue, but I will not address it in this paper.

[12] Note that the pronouns differ in number, but this is as we expect, because (11a) has a plurality inference after all and by the time the pronoun is processed, x is already understood as referring to a plurality. On the other hand, for (11b), x is supposed to be a single pocket, so the pronoun should be singular. We will see how my account of the plurality inference gives rise to these results in detail the next section.

other hand, x in (12b) is restricted to be an atomic pocket. Then by the same reasoning as above, we derive the inference for (12a) that x must refer to a plurality of pockets. Thus, the derivation of the plurality inference in this case is completely parallel to, and as simple as, non-quantified cases like (11a).

3 Discourse referents in dynamic semantics

I will now introduce a simple dynamic semantic theory, in order to formalise the idea I have just sketched (see Appendix A for motivation for this choice of framework). I will augment it with generalised quantifiers in Sect. 5, but the core part of the system will stay the same.

3.1 A primer for dynamic semantics

In dynamic semantics, sentence meanings are modelled as functions over *information states*, which are formal representations of (certain relevant aspects of) discourse contexts. Following Heim (1982), we take information states to be sets of world-assignment pairs (and ingore other aspects of discourse contexts that do not matter for the phenomenon under consideration). Each world-assignment pair is meant to represent a live possibility according to the common ground among the interlocutors at a given point in discourse, so let us call world-assignment pairs *possibilities*. Thus, an information state is a set of possibilities and each possibility is a world-assignment pair.

It will be convenient to be able to refer to just the worlds or just the assignments found in a given information state, so let us use the following functions that bisect the possibilities and discard one of the components.

(13) For any information state c,

a. $\mathsf{W}(c) = \{w | \langle w, a \rangle \in c \text{ for some } a\}$
b. $\mathsf{A}(c) = \{a | \langle w, a \rangle \in c \text{ for some } w\}$

Note that there are versions of dynamic semantics that are simpler than this, e.g. Groenendijk and Stokhof (1991) Dynamic Predicate Logic (DPL), where an information state is a single world-assignment pair, or even just a single assignment without a world, rather than sets thereof. As we will see, for our purposes, it will be useful to explicitly represent sets of world-assignment pairs, and when we discuss presuppositions, this will be crucial, so we will stick to this setup.[13]

The rest of this subsection will introduce the formal details of the dynamic semantic system I will be using. I believe it is largely standard, including the notation, so if the

[13] To put it conversely, the analysis of scalar implicatures *per se* could be reformulated in a point-wise version of dynamic semantics, like DPL, because the core system is entirely distributive (in the sense of Rothschild and Yalcin, 2016), and the only non-distributive part of the system is presuppositions. I will not offer such an alternative version of the theory in this paper, but the necessary formal translation from what is presented below to such an alternative theory should be more or less routine.

reader is familiar with dynamic semantics, they can safely skip the rest of the current subsection.

Recall now that a sentence in natural language has truth-conditions as well as a separate type of meaning related to pronominal anaphora. In the current version of dynamic semantics, this can be thought of as follows: The truth-conditional meaning of a sentence uttered in context c operates on the worlds in $\mathsf{W}(c)$ and its anaphoric meaning operates on the assignments in $\mathsf{A}(c)$. For example, a simple sentence with no quantifiers or connectives have trivial anaphoric meaning and do not change $\mathsf{A}(c)$, but still operates on $\mathsf{W}(c)$, as illustrated in (14). I will use the post-fix notation where $c[\phi]$ is the result of applying ϕ's denotation to information state c, which is dynamic semantics' way of characterizing what happens when an assertion of ϕ is made the discourse context that c represents. We call an application of a sentence denotation to an information state an *update*, and read $c[\phi]$ as 'c updated with ϕ'.

(14) a. $c[\texttt{It is raining in London}] = \{\langle w, a\rangle \in c | \text{it is raining in London in } w\}$
b. $c[\texttt{Paul is a registered taxpayer}]$
$= \{\langle w, a\rangle \in c | \text{Paul is a registered taxpayer in } w\}$

As these representations make clear, these sentences only put constraints on which possible worlds w can remain in the resulting information state.[14] Note, however, that these updates might have indirect consequences on anaphoric possibilities. For instance, one can take an information state c where some assignments in $\mathsf{A}(c)$ map x to Nathan, but all such assignments are paired with worlds where it is sunny in London. When applied to this information state c, the truth-conditional meaning of *It is raining in London* will eliminate all those assignments that map x to Nathan, and as a result, x will be known to be referring to someone or something other than Nathan in the resulting information state.

Assignments are used to enable pronominal anaphora. We assume that pronouns are interpreted as discourse referents, which are modelled as variables, as illustrated below. We will ignore presuppositions in this section, so I will not explicitly represent the information coming from the ϕ-features of these pronouns.

(15) a. $c[\texttt{Nathan emailed him}_y] = \{\langle w, a\rangle \in c | \text{Nathan emailed } a(y) \text{ in } w\}$
b. $c[\texttt{He}_y \texttt{ remembers it}_x] = \{\langle w, a\rangle \in c | a(y) \text{ remembers } a(x) \text{ in } w\}$

By assumption, pronominal anaphora only succeeds if the discourse referent in question has already been introduced in the discourse context. In other words, an information state, which is a slice of a discourse at a particular point in time, specifies which discourse referents are active and accessible at that point. There are several different formal ways of representing this information (see, e.g., Heim, 1982; van Eijck, 2001; Nouwen, 2007), but the following simple idea will do for the purposes of the present paper: Assignments are partial functions from variables to entities, and pronominal anaphora with respect to discourse referent x fails in c when any assignment of $\mathsf{A}(c)$ is undefined for x.

[14] Note that we are ignoring Stalnaker's (1998) idea that an assertion of ϕ will make it commonly known that ϕ has been just uttered (see Appendix A). We could include this bit of information, but it would make no difference with respect to my analysis of plurality inferences.

Discourse referents can be introduced in several distinct ways (Heim, 1982), but the only relevant one for now is via indefinites, as illustrated by (16).

(16) c[`Paul has [a TIN]`^{x}]
$= \{\langle w, a[x \mapsto n]\rangle | \langle w, a\rangle \in c$ and n is Paul's TIN in $w\}$

$a[x \mapsto n]$ is an assignment that is different from a at most in that $x \in \mathrm{dom}(a[x \mapsto n])$ and $a[x \mapsto n](x) = n$. Suppose that $\mathsf{A}(c)$ contains assignments that are not defined for x. Even if that is the case, the output of the update in (16) yields a context where all assignments are defined for x. Contrast this with (14b), which does not introduce a discourse referent x. There, the result of the update will still contain assignments undefined for x (as long as these assignments are paired with worlds where Paul is a registered taxpayer). This accounts for the contrast in (1) that we started out with.

It is convenient to assume that an indefinite is always associated with a new variable with respect to the input information state c, i.e. for each $a \in \mathsf{A}(c)$, $x \notin \mathrm{dom}(a)$, because this prevents information loss. This condition is often called the *Novelty Condition* (cf. Heim, 1982), and I assume it to be a presupposition.[15]

I should remark at this point that in this paper we will not be concerned with sub-sentential compositional semantics. One can build a compositional semantic analysis of sentences like the ones above, and it would ultimately be of interest for my analysis of plurality inferences, especially with respect to the question of what exactly is the mechanism behind introduction of discourse referents, but this is beyond the scope of this single paper. See Groenendijk and Stokhof (1990), Muskens (1996), Brasoveanu (2007), Charlow (2014), among others, for concrete dynamic systems of compositional semantics. For this reason, I will not make an explicit claim as to where exactly the discourse referent is introduced in a sentence containing an indefinite (it could potentially be outside the indefinite), and merely notate which discourse referent is related to a given indefinite with a superscript on it.

For the sake of completeness, we will also introduce some connectives. The negation is interpreted as (17a), which makes reference to the idea of *extension*: Assignment a' is an extension of assignment a, written $a \preceq a'$, iff for each $x \in \mathrm{dom}(a)$, $x \in \mathrm{dom}(a')$ and $a(x) = a'(x)$.

(17) a. c[`not` ϕ]$=\{\langle w, a\rangle \in c |$there is no a' such that $a \preceq a'$ and $\langle w, a'\rangle \in c[\phi]\}$
b. c[ϕ `and` ψ] $= c[\phi][\psi]$

Note that the negation cannot be simply set subtraction, $c - c[\phi]$, in a 'non-eliminative' system like the present one, because if ϕ introduces new discourse referents, set subtraction would be vacuous. We will, however, use set subtraction for computing scalar implicatures, as we will discuss in more detail later.

These are the only connectives I will use. In other words, I will avoid disjunction and conditionals in this paper, because their dynamic analyses are highly controversial. A proper dynamic treatment of disjunction turns out to be elusive, especially with respect to pronominal anaphora (see, for example, Stone, 1992; Krahmer and Muskens,

[15] One could alternatively derive it in terms of an anti-presupposition in reference to the definite version of the sentence, as suggested by Heim (1991). I believe this will eventually be better, but it would introduce unnecessary complications that are orthogonal to my main goal in this paper.

1995; Rothschild, 2017). To make the matter worse, disjunction introduces its own scalar implicature, and generally speaking, sentences containing multiple scalar items involve complications that pose additional theoretical challenges (see, e.g., Chierchia, 2004; Fox, 2007; Romoli, 2012; Franke and Bergen, 2020; Bar-Lev and Fox, 2020 for relevant discussion). Similarly, conditionals are outside the scope this paper. Dynamic theories often offer a dynamic version of material implication, but a material implication analysis of natural language conditionals is known to be inadequate, and it is standardly assumed that a proper analysis of conditionals requires intensional semantics. In addition to the heated debate over which intensional theory of conditionals is more appropriate (see von Fintel, 2011, 2012; Kaufmann and Kaufmann, 2015; Egré and Cozic, 2016 for recent overviews), intensionality will bring in further complications that have to do with *de re* reference, which I cannot deal with in this single paper.

For these reasons, I will reserve discussion of the predictions of my theory with respect to these additional connectives and intensional contexts for future work, but instead of them, I will explore the predictions of the theory with respect to extensional quantificational contexts in Sect. 5.

3.2 Adding plurality

Finally, we need to augment the above simple dynamic semantic theory with plurality, as the main empirical interest of this paper is plurality inferences. Following the standard tradition on plurality (Link, 1983; Schwarzschild, 1996; Landman, 2000; Winter, 2001 among others), we will postulate plural entities, in addition to atomic entities, in our semantic model. From the domain D of a given model, which is the set of atomic entities, we define the domain of entities D_e as the closure of D with the plurality forming operator, $\oplus$. By assumption, $\oplus$ is commutative, associative, and idempotent. $\oplus$ induces a part-whole relation $\sqsubseteq$ in the usual manner: For any $x, y \in D_e$, $x \sqsubseteq y$ iff $x \oplus y = y$. We write $x \sqsubset y$ iff $x \sqsubseteq y$ and $x \neq y$. $\langle D_e, \sqsubseteq\rangle$ is an atomic semi-lattice with the members of D as atoms.

Now we allow discourse referents to refer to plural entities in addition to atomic entities. Recall we assume that a plural noun is semantically number-neutral, while a singular noun is semantically singular. This is represented as in (18) , where I is the interpretation function of the model.

(18) a. $x \in I_w(\texttt{pocket})$ iff $x \in D$ and x is a pocket in w
b. $x \in I_w(\texttt{pockets})$ iff each atomic part of x is a pocket in w

Note that (18b) is number-neutral, because whenever x is an atomic entity, there is only one atomic part of it, namely, x itself. The universal quantification over atomic parts here can be understood as reflecting the inherent distributivity of the plural noun. In this paper, we will not discuss non-distributive predication, but the semantic theory assumed here can be extended to accommodate it (cf. van den Berg, 1996; Brasoveanu, 2007, 2008).

Since we do not deal with sub-sentential compositionality in the present paper, the analysis of nouns themselves is not very important, but what is crucial is that sentences containing these nouns may give rise to different anaphoric possibilities.

Let us consider (19). To simplify the discussion let us assume that *this coat* denotes an atomic coat k in every world in $\mathsf{W}(c)$.

(19) a. c[`This coat has [a pocket]`^{x}]
$$= \left\{ \langle w, a[x \mapsto e] \rangle \,\middle|\, \begin{array}{l} \langle w, a \rangle \in c \text{ and } k \text{ has at least one pocket in } w \\ \text{and } e \text{ is an atomic pocket of } k \text{ in } w \end{array} \right\}$$

b. c[`This coat has pockets`^{x}]
$$= \left\{ \langle w, a[x \mapsto e] \rangle \,\middle|\, \begin{array}{l} \langle w, a \rangle \in c \text{ and } k \text{ has at least one pocket in } w \\ \text{and each part of } e \text{ is a pocket of } k \text{ in } w \end{array} \right\}$$

It is important to understand how exactly the resulting contexts differ, so let us consider a concrete example information state. Let us assume that c is an information state that is ignorant about whether or not k has pockets and is open to the possibility that it has exactly one pocket as well as to the possibility that it has two (but not to the possibility that it has more than two). To stay as simple as possible, let us assume that $\mathsf{W}(c) = \{w_0, w_1, w_2\}$ such that k has no pockets in w_0 and exactly one pocket p in w_1 and exactly two pockets p_L and p_R in w_2. Assignments in c could be anything, but just to have some variation, let us suppose

$$c = \left\{ \begin{array}{l} \langle w_0, a \rangle, \langle w_0, b \rangle, \langle w_0, d \rangle, \\ \langle w_1, a \rangle, \langle w_1, f \rangle, \\ \langle w_2, b \rangle, \langle w_2, f \rangle \end{array} \right\}.$$

There can well be more worlds and more assignments, but I do not want to clutter the exposition too much, so I will work with this toy example. Updating this information state with the above two sentences, we will get the following information states, which I will call c'_s and c'_p respectively.

(20) a. c[`This coat has [a pocket]`^{x}]
$$= \left\{ \begin{array}{l} \langle w_1, a[x \mapsto p] \rangle, \langle w_1, c[x \mapsto p] \rangle, \\ \langle w_2, b[x \mapsto p_L] \rangle, \langle w_2, b[x \mapsto p_R] \rangle, \\ \langle w_2, e[x \mapsto p_L] \rangle, \langle w_2, e[x \mapsto p_R] \rangle \end{array} \right\} = c'_s$$

b. c[`This coat has pockets`^{x}]
$$= \left\{ \begin{array}{l} \langle w_1, a[x \mapsto p] \rangle, \langle w_1, e[x \mapsto p] \rangle, \\ \langle w_2, b[x \mapsto p_L] \rangle, \langle w_2, b[x \mapsto p_R] \rangle, \langle w_2, b[x \mapsto p_L \oplus p_R] \rangle, \\ \langle w_2, e[x \mapsto p_L] \rangle, \langle w_2, e[x \mapsto p_R] \rangle, \langle w_2, e[x \mapsto p_L \oplus p_R] \rangle \end{array} \right\} = c'_p$$

Several remarks are in order. First, neither $\mathsf{W}(c'_s)$ nor $\mathsf{W}(c'_p)$ contains w_0. This is because the truth-conditional meanings of these sentences eliminate any possibility whose world is w_0. Second, notice that $\mathsf{W}(c'_s) = \mathsf{W}(c'_p) = \{w_1, w_2\}$. This reflects the observation mentioned in Sect. 2.2 that pairs of sentences like these are truth-conditionally identical, on the assumption that plural nouns are number-neutral. Third, c'_s and c'_p are nonetheless distinct sets, and their crucial difference comes from the fact that $\mathsf{A}(c'_s) \neq \mathsf{A}(c'_p)$. In particular, each assignment in $\mathsf{A}(c'_s)$ assigns an atomic pocket to x, and while the same assignments can also be found in $\mathsf{A}(c_p)$, $\mathsf{A}(c_p)$ contains additional ones that assign a plurality of pockets, namely $p_L \oplus p_R$, to x. Thus, we have $\mathsf{A}(c'_s) \subset \mathsf{A}(c'_p)$. This is exactly the semantic asymmetry that we will make use of to derive the plurality inference as a scalar implicature, which will be discussed more precisely in the next section.

3.3 Excursus: non-maximality

Note that what is represented above is non-maximal readings of the indefinites in the sense that x is not required to denote a maximal entity with respect to $\sqsubseteq$ in the respective possible worlds that complies with the number marking. More specifically, the maximal reading of the singular indefinite would amount to that k has exactly one pocket, which would rule out possibilities whose world component is w_2. Since *a*-indefinites generally allow for non-maximal readings, the above analysis is fine at least as a possible reading of the sentence. One might also want to derive the maximal reading as a separate reading, which could be achieved by postulating lexical ambiguity in *a*-indefinites (see Brasoveanu, 2007, 2008), or by deriving it as a scalar implicature (see Spector, 2007). This aspect of the semantics of *a*-indefinites is not crucial for my main goal here, so I will leave it open here.

The non-maximality of the plural example is potentially more problematic. This is not obvious in the above example, as there is only one relevant plurality, and the atomic referents will be eventually eliminated due to the plurality inference. However, suppose that the input information state c contains a world w_3 where k has three pockets, p_1, p_2 and p_3. Let us suppose that $\langle w_3, d\rangle \in c$. Then after the update with the plural sentence, the information state c'_p will contain each of the following seven extensions of d, each paired with w_3.

$$\begin{array}{lll} d'_1(x) = p_1 & d'_2(x) = p_2 & d'_3(x) = p_3 \\ d'_4(x) = p_1 \oplus p_2 & d'_5(x) = p_2 \oplus p_3 & d'_6(x) = p_1 \oplus p_3 \\ d'_7(x) = p_1 \oplus p_2 \oplus p_3 & & \end{array}$$

d'_1, d'_2 and d'_3 will eventually be removed by the plurality inference, but we will still have three assignments—d'_4, d'_5, and d'_6—that assign a non-maximal plurality to x, in addition to d'_7, which assigns the maximal entity $p_1 \oplus p_2 \oplus p_3$ in w_3.

It is often remarked in the literature (e.g. Brasoveanu, 2007, 2008) that plural indefinites only allow for maximal readings, based on examples like the following.

(21) This coat has [pockets]^x. They$_x$ are inside.

That is, the second sentence here seems to mean that all the pockets of the coat are inside, rather than at least two of them are inside. This maximality effect is not accounted for by the above semantics. One way to fix it without changing our analysis of plural nouns is to assume that the plural pronoun triggers a maximality effect. That is, it discards all non-maximal values, as in (22).

(22) $c[\texttt{They}_x\ \texttt{are inside}]$

$$= \left\{ \langle w, a\rangle \in c \;\middle|\; \begin{array}{l}\text{each atomic part of } a(x) \text{ is inside in } w \text{ and} \\ \text{for no } \langle w, a'\rangle \in c, a(x) \sqsubset a'(x)\end{array} \right\}$$

It is of course a legitimate question why plural pronouns behave like this. However, I would like to point out that singular pronouns also seem to trigger comparable interpretive effects in similar sentences (cf. Evans, 1980; Heim, 1982). Consider (23).

(23) This coat has [a pocket]^x. It$_x$ is inside.

This sentence has a robust inference that the coat has only one pocket, which is a maximality effect on a par with (22). If this is correct, the semantics of *a*-indefinites we gave above will not account for it by itself. We can fix it by giving a parallel maximal account of the singular pronoun as in (24).

(24) $$c[\texttt{It}_x\ \texttt{is inside}] = \left\{ \langle w, a \rangle \in c \,\middle|\, \begin{array}{l} a(x) \text{ is inside in } w \text{ and} \\ \text{for no } \langle w, a' \rangle \in c,\, a(x) \neq a'(x) \end{array} \right\}$$

Having said this, it is also true that a singular pronoun with an *a*-indefinite antecedent sometimes does not seem to have maximality effects. For instance, (25)does not imply that there is only one supermarket near my flat.

(25) There is [a supermarket]^x near my flat. It$_x$ is open until midnight.

It's not very clear if this is a problem, because the maximal pronoun in (24) can be made compatible with this observation, on the assumption that domain restriction is possible in the first sentence, e.g. to the ones that are 'relevant' in some sense. Rather, the real question is whether singular and plural pronouns and indefinites behave differently. In fact, I am not sure if the plural version of (25) differs from it in this regard. That is, (26) does not seem to imply that all the supermarkets near my flat are open until midnight.

(26) There are [supermarkets]^x near my flat. They$_x$ are open until midnight.

I refrain from making strong empirical claims here about these examples, but as far as I can see, pronouns can be blamed for maximality effects, as suggested above.[16] If so, we do not have to make changes to our non-maximal analyses of singular and plural indefinites.

4 Scalar implicatures with discourse referents

We are now ready to derive the plurality inference of (11a) ("The coat has pockets") as a scalar implicature. According to the analysis given in (19), (11a) and its singular counterpart (11b) ("The coat has a pocket") give rise to the same truth-conditional effects, but different anaphoric potentials. As in (19), we call the resulting information

[16] Note also that this is not particularly surprising, if pronouns and definite descriptions share certain core semantic components, as suggested by their semantic as well as morphosyntactic similarities found across languages (see Postal, 1966; Elbourne, 2005; among many others), although this is not directly reflected in the version of dynamic semantics assumed here. We might ultimately want to pursue a dynamic account that somehow incorporates insights from the analysis of pronouns as disguised definite descriptions, but that is obviously beyond the scope of this paper.

states of these sentences c'_s and c'_p. Their truth-conditional equivalence amounts to the equality $\mathsf{W}(c'_s) = \mathsf{W}(c'_p)$. Crucially, however, whenever there is a world in $\mathsf{W}(c)$ where the coat k has more than one pocket, we are bound to have $\mathsf{A}(c'_s) \subset \mathsf{A}(c'_p)$.[17] This means that the singular version of the sentence is more informative, in the sense to be made clear immediately below. We will then use this semantic symmetry to derive a scalar implicature.

4.1 Informativity in dynamic semantics

Let us first be more precise about the notion of informativity. As remarked in the introduction, most of the current literature on scalar implicature exclusively focuses on informativity with respect to truth-conditions. That is, ϕ is more informative (alt. stronger) than ψ iff whenever ϕ is true, ψ is true but not vice versa. Let us call this notion of informativity *truth-conditional informativity*. In the version of dynamic semantics we are using, it can be formalised as follows.

(27) ϕ is *truth-conditionally more informative* than ψ iff for each information state c, $\mathsf{W}(c[\phi]) \subseteq \mathsf{W}(c[\psi])$, and there is at least one information state c' such that $\mathsf{W}(c'[\psi]) \nsubseteq \mathsf{W}(c'[\phi])$.

According to certain theories of scalar implicature, the relevant notion of informativity is contextually localised, which can be defined as (28).

(28) ϕ is *contextually truth-conditionally more informative* than ψ with respect to information state c iff $\mathsf{W}(c[\phi]) \subset \mathsf{W}(c[\psi])$.

The difference between (27) and (28) is essentially the same as the difference between entailment *simpliciter* vs. contextual entailment, when entailment is understood truth-conditionally.

These notions of informativity are not useful for the phenomenon we are after, because regardless of what the original information state c is, it is guaranteed that $\mathsf{W}(c'_s) = \mathsf{W}(c'_p)$. However, this is not the only notion of informativity, and dynamic semantics with discourse referents lends itself to formally representing notions of informativity that encompass anaphoric information. For instance, we can define notions that are just like (27) and (28) above but are about assignments, which I call *anaphoric informativity*.

(29) ϕ is *anaphorically more informative* than ψ iff for each information state c, $\mathsf{A}(c[\phi]) \subseteq \mathsf{A}(c[\psi])$, and there is at least one information state c' such that $\mathsf{A}(c'[\psi]) \nsubseteq \mathsf{A}(c'[\phi])$.

(30) ϕ is *contextually anaphorically more informative* than ψ with respect to information state c iff $\mathsf{A}(c[\phi]) \subset \mathsf{A}(c[\psi])$.

[17] Conversely, if there is no such world in $\mathsf{W}(c)$, then the two sentences will have contextually equivalent effects. Notice that in that case, the plural version is actually infelicitous. Under the scalar implicature approach, this can be understood as a special case of a general phenomenon that applies to all scalar implicatures, which Magri (2009a, b) discusses in depth (see also Anvari, 2019). We could adopt his proposal to account for the observation, but I omit the details.

Furthermore, we can define notions that refer to both aspects of meaning at the same time:

(31) ϕ is *dynamically more informative* than ψ iff for each information state c, $c[\phi] \subseteq c[\psi]$, and there is at least one information state c' such that $c'[\psi] \nsubseteq c'[\phi]$.

(32) ϕ is *contextually dynamically more informative* than ψ with respect to information state c iff $c[\phi] \subset c[\psi]$.

It should be remarked that ϕ being truth-conditionally more informative than ψ does not imply that ϕ is anaphorically or dynamically more informative.[18] To see this, consider a case where ϕ is truth-conditionally more informative than ψ and introduces a discourse referent, while ψ introduces no discourse referent. More concretely, ϕ = "She has [a baby boy]^x" and ψ = "She is a parent". Since ϕ introduces a new discourse referent, $\mathsf{A}(c[\phi])$ will generally not be a subset of $\mathsf{A}(c[\psi])$, which also means that $c[\phi]$ will not be a subset of $c[\psi]$.

On the other hand, it turns out that for the kind of sentences under consideration in this paper, if ϕ is anaphorically more informative than ψ, then ϕ is truth-conditionally more informative than ψ as well.[19] This means that being anaphorically more informative implies being dynamically more informative. Obviously the converse also holds, so being anaphorically more informative and being dynamically more informative amount to the same thing for the sentences we consider in this paper, although in the general case, they are not equivalent. We could therefore use either notion of informativeness in the following discussion. I will use dynamic informativity as our primary notion of informativity.

4.2 Pragmatic implementation

At this point let us recall Grice's intuition: An utterance of sentence ϕ has a scalar implicature, when there is an alternative sentence ψ that could have been used to mean something more informative. More specifically, the classical Gricean theory derives a scalar implicature by means of the *Maxim of Quantity*.

(33) Maxim of Quantity

a. Make your contribution as informative as is required.
b. Do not make your contribution more informative than is required.

---

[18] I thank Brian Buccola for posing questions about the relations among the three notions of informativity.

[19] Formal proof of this would require an explicit syntax, but here is an informal sketch of the proof. Let ϕ be anaphorically more informative than ψ, and suppose hypothetically that for some c $\mathsf{W}(c[\phi])$ is not a subset of $\mathsf{W}(c[\psi])$. Then there must be some w such that $w \in \mathsf{W}(c[\phi])$ but $w \notin \mathsf{W}(c[\psi])$. Now take an assignment a such that $\langle w, a\rangle \in c$ and construct from a another assignment b that is distinct from a by assigning a value to an irrelevant variable that is not in the domain of a and is not mentioned in ϕ or ψ. Then, we can construct a context $c' = \{\langle w, a\rangle, \langle w', b\rangle\}$ with some $w' \in \mathsf{W}(c[\psi])$ and $w \neq w'$. Note that there must be such w' because the premises imply that $c[\psi]$ is not empty. Since nothing in our system brings two distinct assignments identical and the irrelevant variable never cause an issue, $c'[\phi] = \{\langle w, a'\rangle\}$ where a' is an extension of a and $c'[\psi] = \{\langle w', b'\rangle\}$ where b' is an extension of b. But this contradicts the assumption that ϕ is anaphorically more informative than ψ.

Under this view, it is most natural to understand the relevant notion of informativity as contextual dynamic informativity in (32), because there is no reason to limit the comparison to one particular dimension of meaning, and also because pragmatic reasoning is about possible discourse moves in a particular context, so all one cares about should be contextually localised informativity. Let us see how the plurality inference can be drawn under this theory.

Upon hearing *This coat has pockets*^{x} with a plural indefinite, the hearer notices that the speaker could have uttered *This coat has [a pocket]*^{x} instead, which would have been contextually dynamically more informative.[20] Notice that the reason why the speaker did not use this alternative cannot be because they do not believe it to be true, given that the two sentences are truth-conditionally equivalent. Rather, it must be because the speaker wants x to be able to denote a plural value or values.

At this point, the inference is simply that for at least one $\langle w, a\rangle \in c'_p$, $a(x)$ is a plurality. This is too weak for the plurality inference, which amounts to that for each $\langle w, a\rangle \in c'_p$, $a(x)$ is a plurality. As is well known, implicatures derived by the Maxim of Quantity are generally weaker than scalar implicatures. For instance, take the sentences in (34). I will (tentatively) assume that these sentences introduce no discourse referents so we can zoom in on their truth-conditions (but see Sect. 5, where we revise this assumption).

(34) a. Most of the windows are closed.
b. All of the windows are closed.

The implicature of (34a) predicted via the Maxim of Quantity is that the speaker is not certain that (34b) is true, rather than that they are certain that it is false, because being uncertain about its truth is good enough reason for not asserting it.

Sauerland (2004) called such a weak implicature a *primary implicature*, and proposed that it can be strengthened to a stronger *secondary implicature* with an additional assumption, often called the Opinionatedness Assumption, that the speaker is opinionated about the alternative, i.e. either they are certain that it is true, or they are certain that it is false (see also Horn, 1989; Spector, 2016). Since the derived inference is incompatible with the speaker believing the alternative to be true, they must believe it to be false, which is the scalar implicature.

In order to derive the plurality inference as a scalar implicature under the Gricean theory, therefore, we need an extra assumption comparable to the Opinionatedness Assumption, but it must be distinct from it, simply because the inference is not about the truth/falsity of the alternative. Rather, the necessary assumption is about the values of x, namely, that either the speaker intends x to denote an atomic entity in each possibility, or they want it to denote a plural entity in each possibility in the resulting information state. This assumption, together with the implicature we derived with the Maxim of Quantity, will result in the plurality inference.

But why would one assume that the speaker intends a uniformly singular or uniformly plural discourse referent at all? I concede that I will not be able to provide a completely satisfactory answer to this question here, but I would like to make two

[20] Here, the previous footnote becomes relevant. For the sake of discussion, let us simply assume that the singular and plural versions of the sentence are not contextually equivalent.

remarks. Firstly, unlike the Opinionatedness Assumption, which is about the speaker's epistemic state, what to encode in a discourse referent is completely up to the speaker. That is, the speaker has full control over whether or not to introduce a discourse referent and what to encode in that discourse referent (up to the expressive power of the language being used), as they depend solely on what expressions the speaker choose to use.[21] Secondly, in a language like English, there are broadly two types of nouns, count and mass. Count nouns are generally used for discrete, countable objects and ideas, while mass nouns can be used to describe countable or uncountable objects (Barner and Snedeker, 2005; Bale and Barner, 2009; Chierchia, 1998, 2010; Landman, 2011; Lima, 2018; Link, 1983; Rothstein, 2017; among many others). Thus, perhaps by using a count noun, the speaker signals that countability is relevant, which makes the distinction between singular and plural entities salient. Given that this distinction is salient, the speaker is likely to deem it important, and perhaps it is justifiable that they want their discourse referent to not straddle across this distinction.

At this point, this extra assumption lacks independent empirical evidence, but it does not seem to me to be particularly less plausible or theoretically less natural than the Opinionnatedness Assumption needed for garden-variety scalar implicatures under the Gricean approach to scalar implicatures. It should also be noted that there are other broadly pragmatic accounts of scalar implicatures such as Franke's (2011) Iterated Best Response model and Bergen et al.'s (2016) Rational Speech Act model, which do not require such extra assumptions to derive scalar implicatures. It is not my purpose here to compare different models of scalar implicature computation, but rather to argue that the idea of scalar implicatures with discourse referents is a theoretically legitimate and empirically useful idea, so instead of delving into these different pragmatic theories of scalar implicature, I will give an implementation of the same idea in the grammatical approach below. This will also help us see certain crucial aspects of the present proposal more explicitly.

4.3 Grammatical implementation

According to the grammatical approach to scalar implicatures (Chierchia et al., 2012; among others), scalar implicatures are semantic entailments, rather than pragmatically derived inferences. The currently standard implementation of this idea postulates a phonologically null operator. Following Fox (2007) among others, I will call it *Exh* here, which is standardly defined as (35) in a static semantic framework.

(35) $[\![\mathrm{Exh}(\phi)]\!]^w = 1$ iff $[\![\phi]\!]^w = 1$ and for each excludable alternative ψ to ϕ, $[\![\psi]\!]^w = 0$

Whether a scalar implicature arises depends on whether this operator is present in the structure, as well as on whether 'excludable alternatives' exist.

Several ways of characterising excludable alternatives have been discussed in the literature (see, e.g., Fox, 2007; Spector, 2016), but they are all built on a general theory of alternatives that defines what counts as an alternative to begin with, among which

[21] One might be tempted to speculate that the robustness of plurality inferences mentioned at the end of Sect. 2.2 has to do with this. I leave this idea open here.

excludable ones are identified. Unfortunately, at the present moment, a general theory of alternatives is yet to be constructed (see Katzir, 2007; Fox and Katzir, 2011, for attempts; see also Breheny et al. 2018 for an overview of current open issues), so I will not deal with this issue in this paper, but note that this is a common issue for all theories of scalar implicature, including the Gricean and other pragmatic theories. What is of more interest for us is that different ways of identifying excludable alternatives that are currently entertained in the literature are all based on truth-conditional informativeness. The simplest among them states that ψ is an excludable alternative to ϕ iff ψ is an alternative to ϕ and is truth-conditionally more informative than ϕ. Note that the relevant notation of informativeness is assumed to be blind to contextual information (Fox and Hackl, 2006; Magri, 2009a, b; Fox and Katzir, 2021), so it is to be understood in terms of truth-conditional informativeness *simpliciter*, (27), rather than contextual truth-conditional informativeness, (28). I will not review arguments for the contextual blindness of Exh here, and simply refer the interested reader to the works cited here.

We know that truth-conditional informativity will not be useful for plurality inferences, so we have to use dynamic informativity to define excludable alternatives. In order to do so we first have to make Exh sensitive to discourse referents by dynamicising it. The following analysis captures Grice's core intuition more or less straightforwardly, where $\mathrm{ExclAlt}(\phi, c)$ is the set of excludable alternatives to ϕ with respect to c.

(36) $c[\mathtt{Exh}(\phi)] = \bigcap_{\psi_i \in \mathrm{ExclAlt}(\phi,c)} (c[\phi] - c[\psi_i])$

In most examples we will discuss, there is only one excludable alternative ψ, so the meaning with the scalar implicature will look like $c[\phi]-c[\psi]$, where $c[\phi]$ corresponds to the literal meaning under the Gricean pragmatic approach, and $c[\psi]$ corresponds to what the alternative would have meant had it been used instead in the same context.

Note that this latter meaning is 'negated' in a particular way. That is, all the possibilities that would have arisen by the use of ψ are removed from $c[\phi]$. The Gricean implementation we discussed above also treated scalar implicatures this way, i.e. everything the speaker could have meant by the alternative ψ is excluded from $c[\phi]$. Importantly, also, this way of 'negating' is different from updating $c[\phi]$ with 'not ψ', which, given the rule in (17a), would be:

$$c[\phi][\mathtt{not}\ \psi] = \{\langle w, a\rangle \in c[\phi] | \text{there is no } a' \text{ such that } a \preceq a' \text{ and } \langle w, a'\rangle \in c[\phi][\psi]\}.$$

This would not give us the plurality inference we want. For example, if the relevant excludable alternative to (37) below is (37a), then c[(37)][(37a)], which would be required in computing c[(37)][`not` (37a)], cannot be computed. This is because the discourse referent x in (37a) is required to be new, but it has already been introduced by (37). Furthermore, if the alternative is understood as introducing a new different discourse referent, say y, as in (37b), then c[(37)][`not` (37b)] is bound to be $\emptyset$, because every possibility $\langle w, a\rangle \in c$[(37)] is such that there is at least one pocket in w and that pocket can be a value of y, so every assignment in $\mathsf{A}(c$[(37)]) has an extension in c[(37)][(37b)].

(37) The coat has pockets^x.

a. The coat has [a pocket]^x

b. The coat has [a pocket]^y

Recall also that the semantics of *not* in English cannot be understood in terms of subtraction. That is, the alternative negation rule

$$c[\text{not } \phi] = c - c[\phi]$$

does not work, because whenever ϕ introduces a new discourse referent, the result of this subtraction operation will be simply vacuous.[22] This means that how alternatives are negated in the computation of scalar implicatures is different from how negation in natural language works. Dynamic semantics is useful in making these different notions of 'negation' explicit.

Having dynamicised Exh, let us redefine the notion of excludable alternatives in order to take into consideration the anaphoric dimension of meaning, in addition to the truth-conditional dimension of meaning. We do so by using dynamic informativeness as in (38).[23]

To transpose the static version of Exh as faithfully as possible to the current setting, we assume that the relevant notion of informativeness is blind to contextual information.

(38) ψ is an *excludable alternative* to ϕ iff ψ is an alternative to ϕ and ψ is dynamically more informative than ϕ.

We now have all the ingredients necessary for deriving plurality inferences. Take (37) as an example. (37a) is a dynamically more informative alternative, so it is an excludable alternative to (37). In particular, while $\mathsf{W}(c[(37a)]) = \mathsf{W}(c[(37)])$, we have $\mathsf{A}(c[(37a)]) \subset \mathsf{A}(c[(37)])$ (assuming that $\mathsf{A}(c[(37)])$ contains an assignment that assigns a plurality to x; see fn. 17 for cases where it does not). Then the scalar implicature removes all the assignments, except those that assign pluralities to x.

---

[22] Note that this rule works in an *eliminative* system (in the sense of Rothschild and Yalcin (2016)), where for each sentence ϕ, we have $c[\phi] \subseteq c$. The way we are dealing with discourse referents here makes the system non-eliminative. Although it is not impossible to have discourse referents in an eliminative system, there are certain potential empirical issues with it, so I will not pursue this idea here.

[23] Another definition that is often found in the literature is 'non-weaker alternatives'. In the current setting, it will look like (i).

(i) ψ is an *excludable alternative* to ϕ iff ψ is an alternative to ϕ and ϕ is not dynamically more informative than ψ.

Obviously, this will make more alternatives excludable. We will crucially use this version in Sect. 6. Furthermore, if one wishes, one could incorporate Fox's (2007) notion of *innocent exclusion*, which will look like (ii).

(ii) a. ψ is *innocently excludable* given ϕ and a set A of alternatives to ϕ iff ψ is a member of a maximal consistent set of excludable alternatives given ϕ and A.

b. $S \subseteq A$ is a *consistent set of excludable alternatives given ϕ and A* iff $\bigcap_{\psi_i \in S}(c[\phi] - c[\psi_i]) \neq \emptyset$

Notice that it is crucial that the alternative (37a) introduces the same discourse referent as (37). However, we do not need to restrict relevant alternatives to alternatives with the same discourse referent. That is, even if (37b) counts as an alternative to (37), it is not dynamically more informative, but rather, dynamically independent from (37), so according to the definition of excludable alternatives above, it will not count as an excludable alternative. Furthermore, even if it counted as an excludable alternative (e.g. under one of the alternative definitions in fn. 23), the scalar implicature derived with this alternative would be vacuous, because the assignments in $\mathsf{A}(c[(37b)])$ would be all distinct from the assignments in $\mathsf{A}(c[(38)])$, given the assumption that y has to be new with respect to c.

It should also be pointed out that other types of scalar implicatures can be computed in the same way. Since *some* and disjunction are kind of indefinites themselves, let us look at *most*. The following example has a scalar implicature that not all of the professors commute by bike, for which the alternative in (39b) is crucial.

(39) a. Most of the professors commute by bike.
b. All of the professors commute by bike.

Assuming for now that there is no discourse referents for these sentences (but see the next section), their meanings are analysed as (40).

(40) a. $c[\texttt{Most of the professors commute by bike}]$
$= \{\langle w, a\rangle \in c | \text{the majority of professors commute by bike in } w\}$
b. $c[\texttt{All of the professors commute by bike}]$
$= \{\langle w, a\rangle \in c | \text{all the professors commute by bike in } w\}$

Since the latter sentence is dynamically more informative than the former, *because of the stronger truth-conditional meaning*, the scalar implicature amounts to removing all the possibilities in (40b) from (40a). This amounts to the inference that not all the professors commute by bike. In other words, the present theory can deal with scalar implicatures that arise via a truth-conditionally more informative alternative as in this example, as well as scalar implicatures that arise via an anaphorically more informative alternative as in the previous example. Also, as we will discuss in Sect. 6, the theory makes an interesting prediction for cases involving an alternative that is both truth-conditionally and anaphorically more informative.

I have now presented two different implementations of the idea of scalar implicatures with discourse referents, a pragmatic implementation and a grammatical implementation. I do not think the empirical phenomenon under discussion provides us with a particularly strong argument for or against either of them, but since the grammatical implementation is formally more detailed and also since the theoretical flexibility it makes available will be useful in understanding certain empirical facts discussed in the next section, I will adopt it for the rest of the paper.

4.4 Plural indefinites under negation

We have just waded through all the technicalities of deriving a plurality inference as a scalar implicature in dynamic semantics, but why did we want an account like this?

Recall that one of the reasons is because we want to understand why plural nouns stay number-neutral in negative contexts like in the scope of negation, as in (41).

(41) This coat does not have pockets^x.

This is straightforwardly accounted for by the present analysis as follows. The alternative to this sentence is (42). We understand it under the narrow scope reading of the indefinite.

(42) This coat does not have [a pocket]^x.

Recall how negation is interpreted, namely (17a). This rule is formulated so as to block discourse referents introduced in the scope of negation from being accessed from outside the scope of negation (Heim, 1982). That is, looking from outside the scope of negation, these sentences look as if they have no discourse referents. In fact, pronominal anaphora fails in both cases as demonstrated in (43).

(43) a. This coat does not have pockets^x. #They$_x$ are outside.
b. This coat does not have [a pocket]^x. #It$_x$ is outside.

This means that (41) and (42) trivially have the same anaphoric properties. Furthermore, it is easy to see that they are truth-conditionally equivalent as well, just like their positive counterparts. Consequently, these sentences are dynamically equivalent, and there is absolutely no semantic asymmetry between them. Therefore, no scalar implicature is predicted for (41), capturing the observation that plural nouns are interpreted number-neutrally in the scope of negation.

An obvious prediction of this account is that plural nouns should be interpreted number-neutrally under any operator that similarly shields discourse referents from access from outside. For reasons mentioned at the end of Sect. 3, I will avoid disjunction and conditionals, but it should be noted that empirical facts suggest that disjunction is not an operator of this type. For instance, the anaphora in (44) is possible (Stone, 1992; Rothschild, 2017), and the discourse referent it introduces is interpreted as an atomic individual, as evidenced by the singular number marking on the pronoun in the second sentence of (44).

(44) Irene has [a cat]^x or [a dog]^x. She loves it$_x$ very much.

Consequently, my analysis predicts that the plural version of the first sentence here, (45) below, should have a plurality inference, and this prediction is borne out. That is, (45) has a plurality inference to the effect that the number of pet animals, whether cats or dogs, is more than one.[24]

[24] Things are a bit more complex for cases where not all disjuncts introduce a discourse referent, as in the first sentence of (i). In such a case, simple anaphora to *a cat* does not succeed, as the infelicity of the second sentence indicates.

(i) Irene is either not allowed to have pets, or has [a cat]^x. #It$_x$ is black.

However, as discussed by Rothschild (2017), it is possible to have limited type of anaphora, at least in certain cases, which Rothschild illustrates with the following example.

(45) Irene has cats or dogs.

Again, I will not try to give an explicit account of disjunction here, but examples like above do not seem to pose an issue.

Conditionals similarly involve additional complications. In addition to their inherent intensionality, they are known to license anaphora in modal contexts, as shown in (46).

(46) If Rafael buys [a unicycle]^x, then he will ride it$_x$ every day.

a. ...He will even use it$_x$ for commuting.

b. #...It$_x$ is in his office.

Since anaphora is possible, the prediction of our account is that there will be some plurality inference, but it will not be a full-blown one, as the anaphora is restricted. However, investigating this further would require us to delve into the complicated issue of modal subordination in addition to the issue of *de re* reference, which I would like to set aside in this paper.

Instead of these constructions, we will discuss so-called *quantificational subordination* in the next section with formal details.

4.5 Partial plurality inferences

Another important empirical motivation for the scalar implicature approach to plurality inferences is partial plurality inferences. As explained in informal terms at the end of Sect. 2.3, our explanation for the plurality inference of (47a) below is exactly the same as the simple case with a non-quantificational subject. That is, its plurality inference is derived in relation to its singular counterpart in (47b).

(47) a. Exactly one of these coats has pockets^x. (They$_x$ are inside)
b. Exactly one of these coats has [a pocket]^x (It$_x$ is inside)

We need to wait until the next section to give a proper analysis to the quantifier *exactly one*, but it is already clear that the discourse referent x is accessible in a later discourse, as demonstrated by the continuations in parentheses in (47). Tentatively giving a simple-minded analysis to the quantificational subject, the meanings of these sentences can be represented as follows.

Footnote 24 continued

(ii) Either it's a holiday or [a customer]^x will come in. And if it's not a holiday, he$_x$'ll want to be served.

The plural version of this example would involve a plural pronoun, as in (iii), but its meaning seems to be compatible with the possibility that it is not a holiday and only one customer comes in.

(iii) Either it's a holiday or customers^x will come in. And if it's not a holiday, they$_x$'ll want to be served.

It seems to me that this sentence has a partial plurality inference that if it's not a holiday, more than one customer *might* come in. An analysis of this will require me to say more about the meaning of conditions, so I will simply remark here that (iii) does not seem to pose any obvious threat to the present account, and leave a more precise analysis of (iii) for another occasion.

(48) a. c[`Exactly one of these coats has pockets`^{x}]

$$= \left\{ \langle w, a[x \mapsto p] \rangle \;\middle|\; \begin{array}{l} \langle w, a \rangle \in c \text{ and} \\ \text{exactly one of the coats has one or more pockets in } w \\ \text{and each atomic part of } p \text{ is a pocket} \\ \quad \text{on the unique coat with one or more pockets in } w \end{array} \right\}$$

b. c[`Exactly one of these coats has [a pocket]`^{x}]

$$= \left\{ \langle w, a[x \mapsto p] \rangle \;\middle|\; \begin{array}{l} \langle w, a \rangle \in c \text{ and} \\ \text{exactly one of the coats has one or more pockets in } w \\ \text{and } p \text{ is an atomic pocket} \\ \quad \text{on the unique coat with one or more pockets in } w \end{array} \right\}$$

As in the case of (19) where *this coat* is the subject, the only difference between (48a) and (48b) is whether the discourse referent x can refer to a plurality. Then, (48a) is dynamically less informative than (48b), so by the same reasoning as above, the scalar implicature will be drawn that all possible values of x are pluralities. We will come back to this example in the next section with a more concrete analysis of the quantifier *exactly one*.

Recall also that there is another type of partial plurality inference, exemplified by (49a). As briefly remarked in Sect. 2.2, the partial plurality inference of this sentence is presuppositional in nature and makes this sentence infelicitous when (49b) is felicitous.

(49) a. [Every passenger]^x lost their$_x$ suitcases.
b. [Every passenger]^x lost their$_x$ suitcase.

In this case too, a detailed analysis of the subject quantifier needs to wait until next section, but the analysis of this partial plurality inference does not hinge on it, since we can simply zoom in on how presuppositions work in the present framework. There are different theories of presupposition, but our dynamic semantics is fully compatible with Heim's (1983), so let us adopt it. The intuition behind this theory is that presuppositions are pre-conditions on updates.

In order to apply this idea to the above examples, we obviously cannot sidestep the thorny issue of presupposition projection in quantified sentences. Fortunately for us, it is more or less uncontroversial that universal quantifiers like *every* give rise to *universal presuppositions* (Heim, 1983; Chemla, 2009; Sudo, 2012, 2014; Fox, 2012; but see Beaver, 2001; Beaver and Krahmer, 2001; George, 2008), and it is in fact what Heim's theory predicts for these examples (although this becomes technically apparent only with a concrete analysis of *every*, which I have not introduced yet). Specifically, the presuppositions of the sentences in (49) can be analysed in the following manner using a distinguished information state # representing the state of presupposition failure.

(50) a. c[[`Every passenger`]^x `lost their`$_x$ `suitcases`] $\neq$ # only if for each $\langle w, a \rangle \in c$, $x \notin \mathrm{dom}(a)$ and every passenger in w has one or more suitcases in w.
b. c[[`Every passenger`]^x `lost their`$_x$ `suitcase`] $\neq$ # only if for each $\langle w, a \rangle \in c$, $x \notin \mathrm{dom}(a)$ and every passenger in w has exactly one suitcase in w.

There might be more presuppositions, e.g. the existential presupposition of *every*, but I will omit them here, as they will make no difference. Note that (50b) has a uniqueness presupposition coming from the singular definite *their suitcase*, and so its presuppositional condition as described in (50b) is simply stronger than that of (50a).

Now, we assume that in a situation like this, a scalar inference is drawn in the domain of presuppositions (Heim, 1991, 2011; Percus, 2006; Gajewski and Sharvit, 2012). It is currently actively debated what exactly is the principle behind it (see, e.g., Spector and Sudo, 2017; Marty, 2017; Anvari, 2019) but let us assume the principle in (51b), which is basically a dynamic version of Spector and Sudo's (2017) idea.

(51) a. ϕ is *presuppositionally stronger* than ψ iff for every c such that $c[\psi] = \#$, $c[\phi] = \#$ and there is at least one c' such that $c'[\phi] = \#$ but $c'[\psi] \neq \#$.
b. If ϕ has a presuppositionally stronger alternative ψ, then whenever $c[\psi] \neq \#$, $c[\phi]$ is understood as #.

This principle strengthens the presupposition in (50a) with a scalar inference that its singular counterpart must result in #, which is exactly the partial plurality inference in the presuppositional domain (also, as mentioned in fn. 9, this inference could be strengthened with an auxiliary assumption). One could use any of the principles proposed in the works cited above to obtain essentialy the same results in this case.

According to the current analysis, the inference in question is strictly speaking not a scalar implicature, but it also builds on the intuition that the inference arises in reference to the singular version of the sentence. Note that the derivation of this inference does not actually require dynamic semantics, as discourse referents are not necessary to derive it and one could use a non-dynamic theory of presupposition, but as demonstrated here, it is fully compatible with our dynamic semantics. I will not discuss this type of partial plurality inference any further, as it is orthogonal to the idea of scalar implicatures arising from discourse referents, but as the reader can easily verify, the enriched system to be discussed in the next section will stay compatible with the above account of it.

5 Plurality inferences in quantificational contexts

In this section, we will closely examine plurality inferences triggered in quantificational contexts, and analyse them in an extension of the dynamic semantic system introduced above. This will allow us to give a full account of the examples involving *exactly one* and also make predictions about plurality inferences triggered under other quantifiers.

In classical dynamic semantics (Kamp, 1981; Heim, 1982; Groenendijk and Stokhof, 1991), indefinites are the primary means of introducing discourse referents. Subsequent developments in the 1990s (Van den Berg, 1996; Chierchia, 1995; Kamp and Reyle, 1993; Kanazawa, 1993, 1994) introduced two important ideas: *selective generalised quantifiers* and *quantificational subordination*. Let us discuss them in turn.

When dynamic semantic theories were originally proposed, Heim (1982), in particular, made use of the mechanism of *unselective binding* to account for quantificational behaviour of indefinites in donkey sentences, but it was pointed out by Roth (1987), among others, that an unselective analysis of quantifiers runs into the so-called *Proportion Problem* for donkey anaphora with quantificational DPs (see Kadmon, 1987; Heim, 1990; Chierchia, 1995; Elbourne, 2005; Brasoveanu, 2007). A consequence of the Proportion Problem is that quantificational DPs must be analysed as selective quantifiers, as in static semantics.

Another important realisation made in the 1990s is that quantifiers interact with discourse referents in intricate ways. Firstly, quantifiers can introduce discourse referents, just like indefinites, as demonstrated by (52).

(52) [Most of the MA students]^x chose Semantics II, instead of Phonology II. It seems that they$_x$ liked Semantics I, but they$_x$ don't know that Semantics II is much more difficult.

Here, the pronouns refer to the MA students that chose Semantics II. This anaphora would not be possible were it not for the quantifier in the first sentence, which suggests that the crucial discourse referent is introduced by the quantifier. Furthermore, quantifiers interact with discourse referents introduced by other phrases, as illustrated by (53).

(53) [Every student in this class]^x conducted [an original experiment]^y.

a. ...Then they$_x$ each gave a presentation about it$_y$.
b. #...It$_y$ was conducted online.
c. They$_y$ were all conducted online.

Of particular interest is the continuation in (53a). Here, a singular pronoun is used to refer back to each of the experiments the students conducted. Note that this is not possible in (53b), although it is possible to refer back to the plurality consisting of all these experiments with a plural pronoun, as shown by (53c). The contrast between (53a) and (53b) shows that although the discourse referent y is introduced in the scope of the universal quantifier in the first sentence, it can be referenced in a later discourse, *only if it stands in a particular relation with x there*. The idea is that in the first sentence, x ranges across the atomic students under question one by one, and y holds information about what experiment each of them conducted. Then this distributed information can be referenced in a later discourse, only if x is understood distributively there. If x is not referenced, as in (53b), the atomic values of y are shielded and only the totality of the experiments can be accessed, as in (53c). The anaphoric phenomenon exemplified as in (53a), is often called *quantificational subordination*.

The current literature on dynamic semantics already contains formal ways of dealing with quantificational subordination with selective quantifiers, so I will just borrow one of them here. What is important for our purposes is that when coupled with the idea of scalar implicatures with discourse referents, the resulting system will make a prediction about sentences like (54).

(54) [Every student in this class]^x conducted [original experiments]^y.

That is, although y is under the scope of a universal quantifier, the discourse referent is still accessible, although in limited ways, as shown above. Since (54) introduces a discourse referent y, we should ask if it has a dynamically more informative alternative, and if it does, it should have a scalar implicature. In order to see what exactly the prediction will be, let us first review how selective quantifiers are defined in dynamic semantics and how quantificational subordination is accounted for.

5.1 Selective generalised quantifiers in dynamic semantics

As remarked above, the literature on donkey anaphora has converged on the consensus that quantificational DPs need to be analysed as selective quantifiers. This is not an appropriate place to review the arguments for this claim in detail, so I will just introduce one way of defining selective quantifiers in the version of dynamic semantics we have been assuming thus far.[25]

In Classical Generalised Quantifier Theory (Barwise and Cooper, 1981; Peters and Westerståhl, 2006; among others), a quantificational determiner expresses a relation between two sets, one denoted by the NP and one denoted by the VP. By conservativity, the relation can be seen as a relation between the NP and the intersection of the sets denoted by the NP and VP. These sets are often called the *maxset* and *refset*, respectively.

Selective generalised quantifiers in dynamic semantics will work essentially in the same way, except that the maxset and refset need to be extracted from the dynamic meanings denoted by the NP and VP. Specifically, we will analyse the denotations of NP and VP as dynamic statements, rather than predicates, and a quantificational determiner as an operator over a pair of such dynamic statements. We assume that the syntax generates a representation like (55), where the same variable that appears on the determiner *every* appears in the NP and VP as well.

(55) $\texttt{Every}^x$ $[_{NP}\, x$ $\texttt{linguist}]$ $[_{VP}\, x$ $\texttt{laughed}]$

The dynamic statements denoted by the NP and VP are analysed in the same way as before. These variables essentially behave as pronouns without phi-features.

(56) a. $c[x\ \texttt{linguist}] = \{\langle w, a\rangle \in c | a(x) \text{ is a linguist in } w\}$
b. $c[x\ \texttt{laughed}] = \{\langle w, a\rangle \in c | a(x) \text{ laughed in } w\}$

The quantificational determiner *every* in (55)comes with a variable x, which is assumed to be new. As in the case of indefinites, this condition can be seen as a presupposition, but I won't represent it explicitly here. What the determiner does is that it first extracts the set of values of x that satisfy the NP denotation ϕ and the set of values of x that satisfy both the NP denotation ϕ and the VP denotation ψ, which are the maxset and the refset, respectively, and then it requires that the refset be a subset of the maxset (or equivalently, they are the same set). More specifically, we can analyse the meaning of *every* as follows. For the sake of simplicity, we tentatively analyse a

[25] For reasons I do not know, it is not very common in the relevant literature to use a Heimian framework where information states are sets, as is done here.

quantificational determiner to shield discourse referents in its scope from access from outside, which we will fix later so as to account for quantificational subordination. We will henceforth write $c[x \mapsto e]$ for $\{\langle w, a[x \mapsto e]\rangle | \langle w, a\rangle \in c\}$.

(57) $c[\texttt{Every}^x\ \phi\ \psi]$
$$= \left\{ \langle w, a\rangle \in c \,\middle|\, \begin{array}{l} \{e \in D_e | \text{for some } a', \langle w, a'\rangle \in c[x \mapsto e][\phi]\} \\ \subseteq \{e \in D_e | \text{for some } a', \langle w, a'\rangle \in c[x \mapsto e][\phi][\psi]\} \end{array} \right\}$$

Let us unpack this. Since the resulting set is a subset of c in this representation, the condition on the right is about the world component of each possibility. The condition requires that one set of entities to be a subset of another. What are these sets? Notice that $c[x \mapsto e][\phi]$ is the set of possibilities $\langle w', a'\rangle$ where a' is some extension of $a \in \mathsf{A}(c)$ and $a'(x) = e$ (note that no operator can overwrite the value of x in the current system, so nothing in ϕ can reassign a new value to x). There might be other differences between a and a' in case ϕ introduces more discourse referents, but such differences will not concern us anyway. Among these possibilities $\langle w', a'\rangle$, we are only interested in those whose world component is w. If there is such a possibility, that means that e satisfies the NP meaning ϕ in w. Similarly for $c[x \mapsto e][\phi][\psi]$. Note that ψ is processed after ϕ. This is because ϕ might introduce a discourse referent that can be referenced in ψ, which is the type of anaphoric dependency called donkey anaphora.[26] Thus, the set on the left is the set of entities that satisfy ϕ in w and the set on the right is the set of entities that satisfy both ϕ and ψ in w, i.e. they are the maxset and the refset. Note that $c[x \mapsto e]$ may, and usually does, contain other worlds than w. There is a way to redefine the condition in (57) in terms of the subset of c where the world component is w, but we stick to the formulation in (57) to make the theory compatible with Heim's (1983) theory of presupposition, which requires us to access all relevant worlds at the same time, although presuppositions will play no crucial role in the following discussion.

When applied to the above example, where $\phi = x\,\texttt{linguist}$ and $\psi = x\,\texttt{laughed}$, the updated information state will be:

$$\left\{ \langle w, a\rangle \in c \,\middle|\, \begin{array}{l} \{e \in D_e | e \text{ is a linguist in } w\} \\ \subseteq \{e \in D_e | e \text{ is a linguist and } e \text{ laughed in } w\} \end{array} \right\}.$$

Note in particular:

$$\langle w, a'\rangle \in c[x \mapsto e][x\ \texttt{linguist}] \text{ iff } e \text{ is a linguist in } w$$
$$\langle w, a'\rangle \in c[x \mapsto e][x\ \texttt{linguist}][x\ \texttt{laughed}] \text{ iff } e \text{ is a linguist and laughed in } w$$

It should be easy to see that this amounts to (distributive) universal quantification.

Other quantifiers can be analysed analogously, by changing the relation between the two sets, but we need to be careful with plurality. In particular significant complications will arise with respect to non-distributivity. For *every*, the NP is normally singular

[26] Note that (57) derives the so-called existential reading of donkey anaphora. There are several ways of dealing with the universal reading (Kanazawa, 1993, 1994; Chierchia, 1995; Champollion et al., 2019). As we are not interested in donkey anaphora in this paper, I will not discuss the universal reading here.

(although it is compatible with a plural NP in certain cases, such as *every two weeks*), so e in the two sets will be atomic, forcing the distributive interpretation of the VP, but for quantifiers that are compatible with plural nouns this is not guaranteed. Non-distributive predication in general introduces a lot of complications, which are largely orthogonal to the main purpose of this paper, so I will simplify the discussion below by pretending that all predicates are distributive and cumulative, i.e. whenever a predicate P is true of an entity a, then P is also true of each atomic part of a (distributivity), and whenever P is true of a and b, it is also true of $a \oplus b$ (cumulativity). This will allow us to simplify the semantics of quantifiers considerably because we can treat them as distributive quantifiers, and dispense with an independent distributivity operator. See Van den Berg (1996), Nouwen (2003, 2007) and Brasoveanu (2007, 2008) for extensive discussions of non-distributive predication in dynamic semantics.

It should also be noted that generally, what is counted in a quantified statement is the number of atomic elements rather than the number of distinct entities. For instance, *Exactly three students are French* does not mean that there are exactly three elements in D_e that are French students, which would be true if a, b and $a \oplus b$ are French students, and no other entity is. To achieve the correct interpretation, the quantification needs to be over atomic entities, as in (58). Here, D is the base domain of the model, which is the set of atomic entities.

(58) $c[\texttt{Exactly three}^x\ \phi\ \ \psi]$

$$= \left\{ \langle w, a\rangle \in c \;\middle|\; \begin{array}{l} |\{e \in D | \text{for some } a', \langle w, a'\rangle \in c[x \mapsto e][\phi]\} \\ \cap \{e \in D | \text{for some } a', \langle w, a'\rangle \in c[x \mapsto e][\phi][\psi]\}| = 3 \end{array} \right\}$$

$$= \left\{ \langle w, a\rangle \in c \;\middle|\; |\{e \in D | \text{for some } a', \langle w, a'\rangle \in c[x \mapsto e][\phi][\psi]\}| = 3 \right\}$$

Finally, we will let the quantificational determiner remember the refset. This is to account for discourse anaphora like (52):

(52) [Most of the MA students]^x chose Semantics II, instead of Phonology II. It seems that they$_x$ liked Semantics I, but they$_x$ don't know that Semantics II is much more difficult.

Since we focus on distributive predication, what is referenced is always the supremum of the refset. This is achieved by the following change to the semantics of *every*, where the variable d stores the plurality covering the refset. Here, $\bigoplus S$ is the supremum of S in $\langle D_e, \sqsubseteq\rangle$. Note that thanks to the distributivity of S, whenever S is finite and contains at least one plurality, we have $\bigoplus S \in S$ and it is the unique maximal element in S.

(59) $c[\texttt{Every}^x\ \phi\ \ \psi]$

$$= \left\{ \langle w, a[x \mapsto d]\rangle \;\middle|\; \begin{array}{l} \langle w, a\rangle \in c \text{ and} \\ M = \{e \in D | \text{for some } a', \langle w, a'\rangle \in c[x \mapsto e][\phi]\} \text{ and} \\ R = \{e \in D | \text{for some } a', \langle w, a'\rangle \in c[x \mapsto e][\phi][\psi]\} \\ \text{and } M \subseteq R \\ \text{and } d = \bigoplus R \end{array} \right\}$$

Van den Berg (1996) and Brasoveanu (2007) use one more variable to register the maxset as well, but the maxset is unnecessary for the phenomena we are dealing with, so it will not be represented here. It is easy to generalise (59) to any other quantifier, i.e. all that is needed is to change the fourth line $M \subseteq R$ to the relation that the quantifier expresses in Classical Generalised Quantifier Theory applied to M and R.[27]

However, (59) is clearly insufficient for quantificational subordination, because ϕ and ψ might introduce new discourse referents in them, but according to (59) they will not be accessible later. In order to deal with it, we need to introduce more machinery.

5.2 Quantificational subordination

Let us consider the following example of quantificational subordination again.

(60) [Every student in this class]^x conducted [an original experiment]^y.
They$_x$ each gave a presentation about it$_y$.

The first thing to note is that the most natural reading of the first sentence is that different students conducted different experiments, and this is the reading we are after. On this reading, the second sentence means that each of the students in the class gave a presentation about their own experiment, rather than another student's. This means that in processing the value of y in the second sentence, we need to know which value of x is paired with which value of y.

Van den Berg (1996) proposes an ingenious way to deal with such dependencies between two variables (see also Brasoveanu, 2007, 2008; Nouwen, 2003). His crucial innovation consists in redefining information states as sets of pairs consisting of a world and a *set of assignments*, rather than a single assignment. Each of the assignments in this set will register one pair of a value of x and a value of y in a given world.

Since this change affects the whole system, we will redefine the basic update rules. The changes necessary for non-quantificational cases are more or less mechanical, as in (61). $A(x)$ denotes $\bigoplus\{a(x)|a \in A\}$.[28]

[27] Except that negative quantifiers give rise to a problem, as there might not be any individual in the refset. There are several ways of dealing with this issue. For instance, we could conditionalise the last line above on $\{e \in D | \text{for some } a', \langle w, a' \rangle \in c[x \mapsto e][\phi][\psi]\} \neq \emptyset$, so that when this set is empty, no update on a will be performed. Another possible solution is to decompose negative quantifiers into negation and a positive quantifier (Van den Berg, 1996). A third possibility is to postulate a bottom element in D_e that every predicate is false of Bylinina and Nouwen (2018). We will not deal with negative quantifiers in our examples, so I will not choose among these solutions. See also Van den Berg (1996), Nouwen (2003) and Buccola and Spector (2016) for other issues of negative quantifiers with non-distributive predication, which cannot be discussed here.

[28] As in Brasoveanu (2008) and Dotlačil (2013), but unlike Van den Berg (1996) and Brasoveanu (2007), we assume that an assignment can return a plurality. We ignore presuppositions here, but for instance, c[She$_x$ is a linguist] $\neq$ # iff for each $\langle w, A \rangle \in c$ and $A(x)$ is an atomic female entity in w. It should be noted that in the current system, we will never have an information state with a set A of assignments such that for some $a, a' \in A$, $x \in \text{dom}(a)$ but $x \notin \text{dom}(a')$.

(61) a. $c[\texttt{It is raining}] = \{\langle w, A\rangle \in c | \text{it is raining in } w\}$
b. $c[\texttt{She}_x \texttt{ is a linguist}] = c[x \texttt{ is a linguist}]$
$= \{\langle w, A\rangle \in c | A(x) \text{ is a linguist in } w\}$
c. $c[\texttt{They}_x \texttt{ are linguists}]$
$= \{\langle w, A\rangle \in c | \text{each atomic part of } A(x) \text{ is a linguist in } w\}$
d. $c[\texttt{not } \phi]$
$= \{\langle w, A\rangle \in c | \text{no } a \in A \text{ has an extension } a' \in A' \text{ such that } \langle w, A'\rangle \in c[\phi]\}$
e. $c[\phi \texttt{ and } \psi] = c[\phi][\psi]$

Quantifiers will make use of this additional structure. Since we are only dealing with distributive predicates, we continue to treat all quantifiers as distributive quantifiers. As before, an indefinite introduces a new discourse referent, but each assignment a in the set A of assignments will be updated, as illustrated in (62). Here, x is assumed to be a new variable in c, and $a \preceq_x b$ means that b is different from a at most in that $x \in \text{dom}(b)$.

(62) $c[\texttt{Linguists}^x \texttt{ smiled}]$

$$= \left\{ \langle w, B\rangle \;\middle|\; \begin{array}{l} \langle w, A\rangle \in c \text{ and} \\ \text{for each } a \in A \text{ there is } b \in B \text{ such that } a \preceq_x b \text{ and} \\ \text{for each } b \in B \text{ there is } a \in A \text{ such that } a \preceq_x b \text{ and} \\ \text{each atomic part of } B(x) \text{ is a linguist that smiled in } w \end{array} \right\}$$

For a singular indefinite, $B(x)$ must be an atomic entity, which means that all assignments $b \in B$ must map x to the same atomic entity.

(63) $c[[\texttt{A linguist}]^x \texttt{ smiled}]$

$$= \left\{ \langle w, B\rangle \;\middle|\; \begin{array}{l} \langle w, A\rangle \in c \text{ and} \\ \text{for each } a \in A \text{ there is } b \in B \text{ such that } a \preceq_x b \text{ and} \\ \text{for each } b \in B \text{ there is } a \in A \text{ such that } a \preceq_x b \text{ and} \\ B(x) \text{ is a linguist that smiled in } w \end{array} \right\}$$

In this setup, *every* can be analysed as follows. We write $A[x \mapsto e]$ to mean $\{a[x \mapsto e] | a \in A\}$, and $c[x \mapsto e]$ denotes $\{\langle w, A[x \mapsto e]\rangle | \langle w, A\rangle \in c\}$, and $B \preceq A$ means that for each $b \in B$, there is $a \in A$ such that $b \preceq a$.

(64) $c[\texttt{Every}^x \ \phi \ \psi]$

$$= \left\{ \langle w, A\rangle \;\middle|\; \begin{array}{l} M = \{e \in D | \text{for some } B', \langle w, B'\rangle \in c[x \mapsto e][\phi]\} \text{ and} \\ R = \{e \in D | \text{for some } B', \langle w, B'\rangle \in c[x \mapsto e][\phi][\psi]\} \text{ and} \\ M \subseteq R \text{ and} \\ A \subseteq \bigcup_{e \in R} \{B' | \langle w, B'\rangle \in c[x \mapsto e][\phi][\psi]\} \\ \text{and for some } B \text{ such that } \langle w, B\rangle \in c, B \preceq A \text{ and} \\ \text{and for each } b \in B, \text{ for each } e \in R, \\ \text{there is exactly one } a \in A \text{ such that } a(x) = e \text{ and } b \preceq a \end{array} \right\}$$

The first three lines are essentially the same as before. The fourth line does the crucial distributive quantification. That is, it universally quantifies over the elements of the refset, and collects the results of the updates of ϕ and ψ for each value of x. A needs to be a set of such extended assignments that satisfy the requirements stated in the last three lines: A needs to be an extension of some B in the original information state, and each assignment $b \in B$ must be extended exactly once with respect to each value of the refset. This entails that $A(x) = \bigoplus R$.

This last bit is a little complicated, but part of this complication comes from the fact that we are building the distributivity operator into the quantifier meaning, instead of representing it separately, as well as from the fact that we are sticking to the Heimian setup where each information state is a set of possibilities, against which presuppositions are computed. In order to understand how (64) works, let us look at an example. The variables x and y are assumed to be new in c.

(65) $c[\texttt{Every}^x\ [_{NP}\, x\ \texttt{student}]\ [_{VP}\, x\ \texttt{conducted [an experiment]}^y]]$

$$= \left\{ \langle w, A\rangle \left| \begin{array}{l} M = \{e \in D | e \text{ is a student in } w\} \text{ and} \\ R = \left\{ e \in D \middle| \begin{array}{l} e \text{ is a student that} \\ \quad \text{conducted an experiment in } w \end{array} \right\} \text{ and} \\ M \subseteq R \text{ and} \\ A \subseteq \bigcup_{e \in R} \left\{ B[x \mapsto e][y \mapsto d] \middle| \begin{array}{l} \langle w, B\rangle \in c \text{ and} \\ d \text{ is an experiment in } w \\ \text{and } e \text{ conducted } d \text{ in } w \end{array} \right\} \\ \text{and } B \preceq A \text{ for some } B \text{ such that } \langle w, B\rangle \in c \\ \text{and for each } b \in B, \text{ for each } e \in R, \\ \text{there is exactly one } a \in A \text{ such that } a(x) = e \text{ and } b \preceq a \end{array} \right. \right\}$$

Suppose that there are exactly two students s_1, s_2 in w_s, and s_1 conducted one experiment e_1, s_2 two experiments e_{21} and e_{22}. Suppose also $\langle w_s, \{a\}\rangle \in c$. Then, in the output information state, we will have the following possibilities that contain extensions of $\{a\}$.

$$\left\langle w_s, \left\{ \begin{array}{l} a[x \mapsto s_1][y \mapsto e_1], \\ a[x \mapsto s_2][y \mapsto e_{21}] \end{array} \right\} \right\rangle$$
$$\left\langle w_s, \left\{ \begin{array}{l} a[x \mapsto s_1][y \mapsto e_1], \\ a[x \mapsto s_2][y \mapsto e_{22}] \end{array} \right\} \right\rangle$$

Note importantly that each possibility pairs each student with one experiment that they conducted. This is exactly the information we need to account for quantificational subordination. That is, we analyse the second sentence of the example as follows. To simplify, I will ignore the discourse referent of the indefinite *a presentation about it*.

(66) $c[\texttt{They}_x\ \texttt{each gave a presentation about it}_y]$

$$= \left\{ \langle w, A\rangle \in c \left| \begin{array}{l} \text{for each } e \in D \text{ such that } e \sqsubseteq A(x), \\ \text{the unique } a \in A \text{ such that } a(x) = e \text{ is such} \\ \text{that } a(x) \text{ gave a presentation about } a(y) \text{ in } w \end{array} \right. \right\}$$

Since this part of the example is not of our main concern, I will not dwell on (66). Rather, what is important is that we now have a semantic representation for the first part of the example that has the right amount of anaphoric information.

5.3 Back to partial plurality inferences

Now we are ready to give a full explanation of the partial plurality inference observed with *exactly one*. Applying the general recipe for dynamic selective generalised quantifiers to this quantifier, we obtain (67). Recall that we crucially assume that the plural noun is number-neutral.

(67) c`[Exactly one`^{x} `[`$_{NP}$ x `coat] [`$_{VP}$ x `has pockets`^{y}`]]`

$$= \left\{ \langle w, A\rangle \;\middle|\; \begin{array}{l} M = \{e \in D | e \text{ is a coat in } w\} \text{ and} \\ R = \left\{ e \in D \;\middle|\; \begin{array}{l} e \text{ is a coat} \\ \text{and } e \text{ has one or more pockets in } w \end{array} \right\} \text{ and} \\ |M \cap R| = 1 \text{ (i.e. } |R| = 1) \text{ and} \\ A \subseteq \bigcup_{e \in R} \left\{ B[x \mapsto e][y \mapsto d] \;\middle|\; \begin{array}{c} \langle w, B\rangle \in c \\ \text{and each atomic part of } d \\ \text{is a pocket of } e \text{ in } w \end{array} \right\} \\ \text{and } B \preceq A \text{ for some } B \text{ such that } \langle w, B\rangle \in c \\ \text{and for each } b \in B, \text{ for each } e \in R, \\ \text{there is exactly one } a \in A \text{ such that } a(x) = e \text{ and } b \preceq a \end{array} \right\}$$

Compare this to the version of the sentence with a singular indefinite.

(68) c`[Exactly one`^{x} `[`$_{NP}$ x `coat] [`$_{VP}$ x `has [a pocket]`^{y}`]]`

$$= \left\{ \langle w, A\rangle \;\middle|\; \begin{array}{l} M = \{e \in D | e \text{ is a coat in } w\} \text{ and} \\ R = \left\{ e \in D \;\middle|\; \begin{array}{l} e \text{ is a coat} \\ \text{and } e \text{ has one or more pockets in } w \end{array} \right\} \text{ and} \\ |M \cap R| = 1 \text{ (i.e. } |R| = 1) \text{ and} \\ A \subseteq \bigcup_{e \in R} \left\{ B[x \mapsto e][y \mapsto d] \;\middle|\; \begin{array}{l} \langle w, B\rangle \in c \\ d \in D \text{ is a pocket of } e \text{ in } w \end{array} \right\} \\ \text{and } B \preceq A \text{ for some } B \text{ such that } \langle w, B\rangle \in c \\ \text{and for each } b \in B, \text{ for each } e \in R, \\ \text{there is exactly one } a \in A \text{ such that } a(x) = e \text{ and } b \preceq a \end{array} \right\}$$

The only difference between the two sentences is in the values of y. It can be an atomic entity or a plurality in (67) , but can only be atomic in (68). Let us consider a concrete example information state. Suppose that there are two worlds w_1 and w_2 such that $\langle w_1, \{a\}\rangle, \langle w_1, \{a, b\}\rangle, \langle w_2, \{b\}\rangle \in c$. Suppose further that in w_1, a coat k_1 has exactly one pocket p_1 and no other coat has any pockets. In w_2, a coat k_2 has two pockets, p_{21} and p_{22} and no other coat has any pockets. Then in the output information state of (68), we have the following possibilities.

$$\begin{array}{l} \langle w_1, \{a[x \mapsto s_1][y \mapsto p_1]\}\rangle \\ \langle w_1, \{a[x \mapsto s_1][y \mapsto p_1], b[x \mapsto s_1][y \mapsto p_1]\}\rangle \\ \langle w_2, \{b[x \mapsto s_2][y \mapsto p_{21}]\}\rangle \\ \langle w_2, \{b[x \mapsto s_2][y \mapsto p_{22}]\}\rangle \end{array}$$

The first possibility is an extension of $\langle w_1, \{a\}\rangle$, the second possibility is an extension of $\langle w_1, \{a, b\}\rangle$, and the last two are extensions of $\langle w_2, \{b\}\rangle$. In the output information state of (67) , on the other hand, there will be more possibilities, because y can be mapped to a plurality. From the same three possibilities, we will get the following five.

$$\langle w_1, \{a[x \mapsto s_1][y \mapsto p_1]\}\rangle$$
$$\langle w_1, \{a[x \mapsto s_1][y \mapsto p_1], b[x \mapsto s_1][y \mapsto p_1]\}\rangle$$
$$\langle w_2, \{b[x \mapsto s_2][y \mapsto p_{21}]\}\rangle$$
$$\langle w_2, \{b[x \mapsto s_2][y \mapsto p_{22}]\}\rangle$$
$$\langle w_2, \{b[x \mapsto s_2][y \mapsto p_{21} \oplus p_{22}]\}\rangle$$

Since there are more possibilities in the second case, the plural version of the sentence is dynamically less informative. As before, the scalar implicature amounts to subtracting all the possibilities that are covered in the first case. Then we are left with the possibilities where y is mapped to a plurality. Note that the possibilities whose world component is w_1 will be eliminated, because y will never be mapped to a plurality there. This means that it will be entailed that the unique coat that has one or more pockets has multiple pockets, which is the plurality inference.

Recall that among the previous theories of plurality inferences, only Spector's (2007) higher-order implicature theory can explain partial plurality inferences of this type (cf. Ivlieva, 2014). The analysis put forward here is conceptually simpler in that no particular assumptions about alternatives are necessary. That is, the anaphoric properties of quantifiers are given independent empirical evidence, and since discourse referents carry information, it is natural to expect them to give rise to scalar implicatures. On the other hand, Spector's (2007) analysis makes crucial use of alternatives that already have scalar implicatures. I would like to make it explicit again that I have nothing against the idea of alternatives with implicatures *per se*. Rather, my main point here is that such an assumption is simply unnecessary to derive plurality inferences, partial or full, once we recognised discourse referents. Furthermore, there is another respect in which our analysis diverges from Spector, to which we now turn.

5.4 Plurality inferences under universal quantifiers

The present account makes a prediction about sentences like (69) (fn. 7).

(69) Every coat has pockets.

The predicted plurality inference for this example will be partial, i.e. at least one coat has multiple pockets. Let us see why. We analyse the literal meaning of this sentence as follows.

(70) c[`Every`^{x} `[`$_{NP}$ x `coat] [`$_{VP}$ x `has pockets`^{y}`]]`

$$= \left\{ \langle w, A \rangle \;\middle|\; \begin{array}{l} M = \{e \in D | e \text{ is a coat in } w\} \text{ and} \\ R = \left\{ e \in D \;\middle|\; \begin{array}{l} e \text{ is a coat} \\ \text{and } e \text{ has one or more pockets in } w \end{array} \right\} \text{ and} \\ M \subseteq R \text{ and} \\ A \subseteq \bigcup_{e \in R} \left\{ B[x \mapsto e][y \mapsto d] \;\middle|\; \begin{array}{c} \langle w, B \rangle \in c \\ \text{and each atomic part of } d \\ \text{is a pocket of } e \text{ in } w \end{array} \right\} \\ \text{and } B \preceq A \text{ for some } B \text{ such that } \langle w, B \rangle \in c \\ \text{and for each } b \in B, \text{ for each } e \in R, \\ \text{there is exactly one } a \in A \text{ such that } a(x) = e \text{ and } b \preceq a \end{array} \right\}$$

The singular version of this sentence means (71) .

(71) c[`Every`^{x} `[`$_{NP}$ x `coat] [`$_{VP}$ x `has [a pocket]`^{y}`]]`

$$= \left\{ \langle w, A \rangle \;\middle|\; \begin{array}{l} M = \{e \in D | e \text{ is a coat in } w\} \text{ and} \\ R = \left\{ e \in D \;\middle|\; \begin{array}{l} e \text{ is a coat} \\ \text{and } e \text{ has one or more pockets in } w \end{array} \right\} \text{ and} \\ M \subseteq R \text{ and} \\ A \subseteq \bigcup_{e \in R} \left\{ B[x \mapsto e][y \mapsto d] \;\middle|\; \begin{array}{l} \langle w, B \rangle \in c \\ d \in D \text{ is a pocket of } e \text{ in } w \end{array} \right\} \\ \text{and } B \preceq A \text{ for some } B \text{ such that } \langle w, B \rangle \in c \\ \text{and for each } b \in B, \text{ for each } e \in R, \\ \text{there is exactly one } a \in A \text{ such that } a(x) = e \text{ and } b \preceq a \end{array} \right\}$$

It is easy to see that (71) is dynamically more informative than (70), as in the previous example. To see this more concretely, let us consider an example information state. Suppose that $\langle w_1, \{a\}\rangle$, $\langle w_2, \{b\}\rangle$, $\langle w_2, \{a, b\}\rangle \in c$, and that in both of these worlds, there are three coats, k_1, k_2, and k_3. In w_1, k_1 has one pocket p_1, k_2 has two pockets p_{21} and p_{22}, and k_3 has two pockets p_{31} and p_{32}. In w_2, each of them has exactly one pocket, p_1, p_2 and p_3. Then in the output information state in (71), we have the following possibilities that come from these three possibilities in c.

$$\left\langle w_1, \left\{ \begin{array}{l} a[x \mapsto k_1][y \mapsto p_1], \\ a[x \mapsto k_2][y \mapsto p_{21}], \\ a[x \mapsto k_3][y \mapsto p_{31}] \end{array} \right\} \right\rangle \left\langle w_2, \left\{ \begin{array}{l} b[x \mapsto k_1][y \mapsto p_1], \\ b[x \mapsto k_2][y \mapsto p_2], \\ b[x \mapsto k_3][y \mapsto p_3] \end{array} \right\} \right\rangle$$
$$\left\langle w_1, \left\{ \begin{array}{l} a[x \mapsto k_1][y \mapsto p_1], \\ a[x \mapsto k_2][y \mapsto p_{22}], \\ a[x \mapsto k_3][y \mapsto p_{31}] \end{array} \right\} \right\rangle \left\langle w_2, \left\{ \begin{array}{l} a[x \mapsto k_1][y \mapsto p_1], b[x \mapsto k_1][y \mapsto p_1], \\ a[x \mapsto k_2][y \mapsto p_2], b[x \mapsto k_2][y \mapsto p_2], \\ a[x \mapsto k_3][y \mapsto p_3], b[x \mapsto k_3][y \mapsto p_3] \end{array} \right\} \right\rangle$$
$$\left\langle w_1, \left\{ \begin{array}{l} a[x \mapsto k_1][y \mapsto p_1], \\ a[x \mapsto k_2][y \mapsto p_{21}], \\ a[x \mapsto k_3][y \mapsto p_{32}] \end{array} \right\} \right\rangle$$
$$\left\langle w_1, \left\{ \begin{array}{l} a[x \mapsto k_1][y \mapsto p_1], \\ a[x \mapsto k_2][y \mapsto p_{22}], \\ a[x \mapsto k_3][y \mapsto p_{32}] \end{array} \right\} \right\rangle$$

In the output information state of (70), on the other hand, there are more possibilities, because y can be mapped to a plurality. Thus, in addition to the possibilities above, we also have.

$$\left\langle w_1, \left\{\begin{array}{l} a[x \mapsto k_1][y \mapsto p_1], \\ a[x \mapsto k_2][y \mapsto p_{21\oplus 22}], \\ a[x \mapsto k_3][y \mapsto p_{31}] \end{array}\right\}\right\rangle \left\langle w_1, \left\{\begin{array}{l} a[x \mapsto k_1][y \mapsto p_1], \\ a[x \mapsto k_2][y \mapsto p_{21\oplus 22}], \\ a[x \mapsto k_3][y \mapsto p_{32}] \end{array}\right\}\right\rangle$$

$$\left\langle w_1, \left\{\begin{array}{l} a[x \mapsto k_1][y \mapsto p_1], \\ a[x \mapsto k_2][y \mapsto p_{21}], \\ a[x \mapsto k_3][y \mapsto p_{31} \oplus p_{32}] \end{array}\right\}\right\rangle \left\langle w_1, \left\{\begin{array}{l} a[x \mapsto k_1][y \mapsto p_1], \\ a[x \mapsto k_2][y \mapsto p_{22}], \\ a[x \mapsto k_3][y \mapsto p_{31} \oplus p_{32}] \end{array}\right\}\right\rangle$$

$$\left\langle w_1, \left\{\begin{array}{l} a[x \mapsto k_1][y \mapsto p_1], \\ a[x \mapsto k_2][y \mapsto p_{21\oplus 22}], \\ a[x \mapsto k_3][y \mapsto p_{31} \oplus p_{32}] \end{array}\right\}\right\rangle$$

And these additional ones are the ones that will remain after the scalar implicature is computed. Of importance here is the fact that in w_1, k_1 only has one pocket, but w_1 will be in the resulting information state of (70), because w_1 can be mapped with a set of assignments where k_2 and/or k_3 are paired with a plurality of pockets. Generally a world w will remain after the update in (70), if at least one coat has multiple pockets there, because in such a case it can be paired with a set of assignments that its singular counterpart cannot give rise to. Consequently, the plurality inference is partial.

It is empirically desirable that this partial plurality inference can be derived, but it should be pointed out that the sentence seems to also have a stronger, fully plural reading that every coat has multiple pockets (see Stateva et al., 2016 for experimental evidence for this ambiguity). Under the grammatical implementation of the current analysis, this stronger reading can be derived as an embedded implicature that is derived by applying Exh at the VP-level in the scope of the universal quantifier. Since this type of ambiguity is also observed with other scalar implicatures, e.g. (72), it does not seem to be particularly problematic.

(72) Every student has done most of the reading.

However, there is a remaining question about how robust embedded implicatures are in these sentences. It is not necessarily the case that the embedded implicature of (72) is as robust as that of (69), which is a complication that exists in addition to the diversity in robustness observed across scalar items mentioned in Sect. 2.2. This requires more empirical research, and is left unaddressed here.

It should also be noted that Spector's (2007) version of the scalar implicature approach predicts the stronger, fully plural reading by default, and in order to account for the weaker, partial plurality inference, he introduces an additional assumption, namely that *every* optionally has *some* as an alternative. I do not have qualms about Spector's assumption, but it is obvious that our analysis is conceptually much simpler.

6 Disjunction under a universal quantifier

The present analysis makes an interesting prediction about sentences like (73).

(73) Every applicant speaks French or German.

This sentence involves a disjunction that we have been eschewing for reasons mentioned at the end of Sect. 3.1, but a DP-level disjunction like this behaves essentially like an existential quantifier with the disjuncts as its domain of quantification, so we can deal with it in the dynamic semantic system at hand (cf. Roth and Partee, 1982; Schlenker, 2006).

Furthermore, our prediction for (73) has some bearing on Crnič's et al. (2015) observation about them. They observe that (72) has a reading whose scalar implicature is that at least one applicant speaks French and at least one applicant speaks German, without implying that not everyone speaks French or that not every one speaks German. That is, the sentence seems to be acceptable and true with respect to a context where every applicant speaks French and a subset of them speak German, for example.

As Crnič et al. point out, furthermore, this scalar implicature does not follow from the standard way of computing scalar implicatures (see also Bar-Lev and Fox, 2020). Here is why. It is often assumed that a disjunctive sentence has the disjuncts as alternatives, in addition to the conjunctive version of the sentence (Sauerland, 2004; Spector, 2016), so the alternatives for (73) are (74).

(74) a. Every applicant speaks French.
b. Every applicant speaks German.
c. Every applicant speaks French and German.

Each of these alternatives is excludable, but negating (74a) and (74b) will conflict with the reading we are after (although they might be appropriate for a different reading of the sentence; see below). If they are not alternatives, on the other hand, the resulting inference will be too weak, as it will be compatible with no applicant speaking German, for example.

In order to derive the relevant reading, Crnič et al. make use of embedded implicatures and also crucially assume that certain alternatives can optionally be ignored in the computation of scalar implicatures, a process called *pruning*. More recently, Bar-Lev and Fox (2020) put forward a different analysis that makes use of what they call *innocent inclusion*, but they also need pruning to derive the reading under consideration.[29] While I will not directly argue against these analyses here, I will demonstrate that by including discourse referents, this reading can easily be derived without recourse to pruning.

[29] Bar-Lev and Fox (2020) do not explicitly discuss this point, so let me add some details. In order to derive the reading in terms of innocent inclusion, they crucially assume that *every* has *some* as an alternative (cf. Spector's (2007) analysis of the partial plurality inference mentioned above). Then in addition to the alternatives in (74), we will also have the following as alternatives.

(i) a. Some applicant speaks French.
b. Some applicant speaks German.
c. Some applicant speaks French and German.

The crucial effect of including these alternatives is that (ia) and (ib) make (74a) and (74b) non-excludable, as they assume innocent exclusion as their notion of excludability and (ia) and (ib) bring in symmetry to block exclusion of (74a) and (74b). However, notice that (ic) is still excludable so we will derive the scalar implicature that no applicant speaks both French and German, but it conflicts with the reading we are after. Thus, to derive this reading, they need to assume that (ic) can be pruned, without pruning (ia) and (ib).

6.1 Disjunction and maximality

Let us first given an analysis to the DP disjunction *French or German*. Although I cannot give a general analysis of disjunction here, we can regard this DP disjunction as an existential quantifier. First, note that it introduces a discourse referent, as shown in (75).

(75) Daniel speaks [French or German]^y. He learned it$_y$ at high school.

A disjunction in an unembedded context like this generally has two types of implicature: an *ignorance implicature* that the speaker is not sure if Daniel speaks French and the speaker is not sure if Daniel speaks German; and an *exclusivity implicature* that the speaker is certain that Daniel does not speak both of these languages. In the following discussion the ignorance implicature will not play a big role, as the reading of (73) we are interested in does not have an ignorance implicature. So let us focus on the exclusivity implicature.

The exclusivity implicature arises from the conjunctive alternative in (76) .

(76) Daniel speaks [French and German]^y. (He learned them$_y$ at high school.)

As shown by the continuation in parentheses, this conjunctive DP introduces a discourse referent. The following simple analysis is enough to account for its anaphoric properties.

(77) c[`Daniel speaks [French and German]`^{y}

$$= \left\{ \langle w, A[y \mapsto \text{French} \oplus \text{German}]\rangle \;\middle|\; \begin{array}{l} \langle w, A\rangle \in c \text{ and Daniel speaks} \\ \text{both French and German in } w \end{array} \right\}$$

If we analysed the disjunction in (75) as an indefinite, as in (78) , however, we would not be able to derive the exclusivity inference.

(78) c[`Daniel speaks [French or German]`^{y}]

$$= \left\{ \langle w, B\rangle \;\middle|\; \begin{array}{l} \text{for some } \langle w, A\rangle \in c, \\ \text{for each } a \in A \text{ there is } b \in B \text{ such that } a \preceq_y b \text{ and} \\ \text{for each } b \in B \text{ there is } a \in A \text{ such that } a \preceq_y b \text{ and} \\ \text{each atomic part of } B(y) \text{ is French or German} \\ \quad \text{and Daniel speaks each atomic part of } B(y) \text{ in } w \end{array} \right\}$$

Given this analysis, the alternative in (77) is dynamically more informative, so it should give rise to a scalar implicature, but it is actually not strong enough to give us the exclusivity inference. This is due to the non-maximality of indefinites discussed in Sect. 3.3 which results in too many possibilities. Essentially, the alternative in (77) is too specific to be able to exclude all of the ones that we want to exclude. Concretely, suppose that $\langle w_{fg}, \{a\}\rangle \in c$ such that Daniel speaks both French and German in w_{fg}. Then the update with the conjunctive alternative (77) will extend this possibility to:

$$\langle w_{fg}, \{a[y \mapsto \text{French} \oplus \text{German}]\}\rangle.$$

On the other hand, (77) will yield two more possibilities, because indefinites are non-maximal.

$$\langle w_{fg}, \{a[y \mapsto \text{French} \oplus \text{German}]\}\rangle$$
$$\langle w_{fg}, \{a[y \mapsto \text{French}]\}\rangle$$
$$\langle w_{fg}, \{a[y \mapsto \text{German}]\}\rangle$$

The scalar implicature will remove the first possibility, but the latter two will remain. This means that after the computation of the scalar implicature, we will still have w_{fg} in the resulting information state, so we are not excluding the possibility that Daniel speaks both French and German. Rather, we are just guaranteeing that y doesn't refer to French $\oplus$ German. But what we want to derive as the exclusivity inference is that he certainly doesn't speak both, so we want to get rid of any possibility whose world component is w_{fg}.

In order to derive the exclusivity inference, we need to assume that the DP disjunction encodes maximality so that it only introduces the first possibility above. This is achieved by the following analysis.

(79) c[`Daniel speaks [French or German]`^{y}]

$$= \left\{ \langle w, B \rangle \;\middle|\; \begin{array}{l} \text{for some } \langle w, A\rangle \in c, \\ \text{for each } a \in A \text{ there is } b \in B \text{ such that } a \preceq_y b \text{ and} \\ \text{for each } b \in B \text{ there is } a \in A \text{ such that } a \preceq_y b \text{ and} \\ \text{Daniel speaks French and/or German in } w \text{ and} \\ \text{if Daniel speaks both of them in } w, \\ \quad \text{for each } b \in B,\, b(y) \text{ is French}\oplus\text{German, and} \\ \text{if Daniel speaks French but not German in } w, \\ \quad \text{for each } b \in B,\, b(y) \text{ is French, and} \\ \text{if Daniel speaks German but not French in } w, \\ \quad \text{for each } b \in B,\, b(y) \text{ is German} \end{array} \right\}$$

This meaning encodes maximality in the sense that for each world where Daniel speaks at least one of the languages, y stores the maximal entity among French, German, and French$\oplus$German that Daniel speaks in that world. Generally, selective generalised quantifiers have maximality in the present system, as they require the refset to be covered by the values of the variable they are associated with, so this essentially means that DP disjunction is analysed as an existential quantifier, rather than as an (non-maximal) indefinite.[30]

[30] Jeremy Kuhn (p.c.) pointed out to me an important observation that the 'exact' readings of bare numerals give rise to a similar problem, if bare plurals are non-maximal indefinites. Specifically, let us tentatively analyse *two languages* to be a non-maximal indefinite with 'exact' semantics. Then an update with *Daniel speaks two*^{y} *languages*, for example, will result in possibilities $\langle w, A[y \mapsto l_1 \oplus l_2]\rangle$, where l_1 and l_2 are languages that Daniel speaks in w. Due to non-maximality, in the general case, some of these possibilities may be formed with possible worlds where Daniel speaks more than two languages. Consider now the alternative *Daniel speaks three*^{y} *languages*. With the non-maximal indefinite analysis for *three languages*, an update with this alternative will result in possibilities $\langle w, A[y \mapsto l_1 \oplus l_2 \oplus l_3]\rangle$ such that l_1, l_2, and l_3 are languages that Daniel speaks in w. Then the results of the two updates have nothing in common, and therefore the present theory predicts no scalar implicature. Furthermore, assigning 'at least' semantics to numerals will not help here. That is, although in this case, there will be a non-trivial scalar implicature, it

6.2 Distributivity inference

Let us now see what the theory predicts for (73). By combining our analysis of *every* and the above analysis of DP disjunction, we obtain (80).

(80) c[`Every`^{x} [$_{NP}$ x `applicant`] [$_{VP}$ x `speaks` [`French or German`]^y]]

$$= \left\{ \langle w, A\rangle \;\middle|\; \begin{array}{l} M = \{e \in D \mid e \text{ is an applicant in } w\} \text{ and} \\ R = \left\{ e \in D \;\middle|\; \begin{array}{c} e \text{ is an applicant and } e \text{ speaks} \\ \text{French and/or German in } w \end{array} \right\} \text{ and} \\ M \subseteq R \text{ and} \\ A \subseteq \bigcup_{e \in R} \left\{ B[x \mapsto e][y \mapsto d] \;\middle|\; \begin{array}{l} \langle w, B\rangle \in c \\ \text{if } e \text{ speaks French and German in } w, \\ \quad d \text{ is French}\oplus\text{German} \\ \text{if } e \text{ speaks French but not German in } w, \\ \quad d \text{ is French} \\ \text{if } e \text{ speaks German but not French in } w, \\ \quad d \text{ is German} \end{array} \right\} \\ \text{and } B \preceq A \text{ for some } B \text{ such that } \langle w, B\rangle \in c \\ \text{and for each } b \in B, \text{ for each } e \in R, \\ \text{there is exactly one } a \in A \text{ such that } a(x) = e \text{ and } b \preceq a \end{array} \right\}$$

The conjunctive alternative (74c) will mean (81).

(81) c[`Every`^{x} [$_{NP}$ x `applicant`] [$_{VP}$ x `speaks` [`French and German`]^y]]

$$= \left\{ \langle w, A\rangle \;\middle|\; \begin{array}{l} M = \{e \in D \mid e \text{ is an applicant in } w\} \text{ and} \\ R = \left\{ e \in D \;\middle|\; \begin{array}{c} e \text{ is an applicant and } e \text{ speaks} \\ \text{French and German in } w \end{array} \right\} \text{ and} \\ M \subseteq R \text{ and} \\ A \subseteq \bigcup_{e \in R} \left\{ B[x \mapsto e][y \mapsto d] \;\middle|\; \begin{array}{l} \langle w, B\rangle \in c \\ d \text{ is French}\oplus\text{German} \\ \text{and } e \text{ speaks } d \text{ in } w \end{array} \right\} \\ \text{and } B \preceq A \text{ for some } B \text{ such that } \langle w, B\rangle \in c \\ \text{and for each } b \in B, \text{ for each } e \in R, \\ \text{there is exactly one } a \in A \text{ such that } a(x) = e \text{ and } b \preceq a \end{array} \right\}$$

This is dynamically more informative than (80), so it will give rise to a scalar implicature. This scalar implicature will remove $\langle w, A[y \mapsto \text{French} \oplus \text{German}]\rangle$ where every applicant speaks both languages in w. Concretely, whenever $\langle w_{FG}, A\rangle \in c$ such that every applicant speaks both languages in w_{FG}, (81) will give rise to the possibility $\langle w_{FG}, A'\rangle$ where for each $a' \in A'$, $a'(y) = \text{French} \oplus \text{German}$. This is the

Footnote 30 continued
will not result in the desired entailment that Daniel does not speak no more than two languages, because in the result of the update with *Daniel speaks two*^{y} *languages*, we can still have possibilities that look like $\langle w, A[y \mapsto l_1 \oplus l_2]\rangle$ with w being a possible world in which Daniel speaks three or more languages, and such possibilities won't be eliminated by the scalar implicature, due to the way the assignments are updated. A solution to the present issue comes about if we assume that bare numerals, just like disjunction, encode maximality (see Buccola and Spector, 2016 for related discussion). Assuming the 'at least' semantics for bare numerals, an update with *Daniel speaks two*^{y} *languages* will only result in possibilities $\langle w, A[y \mapsto l_1 \oplus \ldots l_n]\rangle$, where $l_1 \oplus \cdots \oplus l_n$ ($n \geq 2$) is the maximal plurality of languages that Daniel speaks in w. Then by eliminating all the possibilities that the alternative with *three* in place of *two* gives rise to, all the possibilities except for those where $n = 2$ will be eliminated, which is the 'exact' reading. I thank Benjamin Spector (p.c.) for a discussion on this point.

only type of possibilities that can be found in the output information state of (81). With respect to the same $\langle w_{FG}, A\rangle \in c$ (80) will give rise to exactly the same possibilities as its conjunctive alternative, so after the scalar implicature is computed, there will be no world like w_{FG} where every applicant speaks both languages in the final information state.

We have two more alternatives. The alternative in (74a), which is without the second disjunct, is analyse as (82). As in the case of a conjoined DP, we assume *French* introduces a discourse referent, which is justifiable as it feeds pronominal anaphora.

(82) c[Every^x [$_{NP}$ x applicant] [$_{VP}$ x speaks French^y]]

$$= \left\{ \langle w, A\rangle \left| \begin{array}{l} M = \{e \in D | e \text{ is an applicant in } w\} \text{ and} \\ R = \left\{ e \in D \left| \begin{array}{l} e \text{ is an applicant and } e \text{ speaks} \\ \text{French in } w \end{array} \right. \right\} \text{ and} \\ M \subseteq R \text{ and} \\ A \subseteq \bigcup_{e \in R} \left\{ B[x \mapsto e][y \mapsto \text{French}] \left| \begin{array}{l} \langle w, B\rangle \in c \\ e \text{ speaks French in } w \end{array} \right. \right\} \\ \text{and } B \preceq A \text{ for some } B \text{ such that } \langle w, B\rangle \in c \\ \text{and for each } b \in B, \text{ for each } e \in R, \\ \text{there is exactly one } a \in A \text{ such that } a(x) = e \text{ and } b \preceq a \end{array} \right. \right\}$$

This is actually neither dynamically more informative nor dynamically less informative than (80). This is because it can introduce possibilities that cannot be introduced by (80). That is, with respect to if $\langle w_{FG}, A\rangle \in c$ such that every applicant speaks both languages in w_{FG}, (82) will give rise to the possibility $\langle w_{FG}, A'\rangle$ where for each $a' \in A'$, $a'(y) =$ French, which is not possible for (80) due to maximality. Furthermore, the output information state of the latter can include possibilities that do not exist in the output information state of (82), e.g. $\langle w_{fg}, A'\rangle$ where some applicants speak only French and the others only speak German in w_{fg}. Therefore, the two sentences are dynamically independent.

Let us assume crucially that a scalar implicature can be computed from such a dynamically independent alternative. As noted in fn. 23, this is achieved by changing the definition of excludable alternatives from those that are dynamically more informative to those that are not dynamically less informative.

Now, let us consider which possibilities will be removed by the scalar implicature (82) triggers. In both output information states we can find the following three kinds of worlds: worlds w_F where everyone speaks French, worlds w_{Fg} where everyone speaks French but only some speak German, and w_{FG} where everyone speaks both. Since all the possibilities with w_{FG} will be removed by the scalar implicature of the conjunctive alternative anyway, we can ignore them here. If $\langle w_F, A\rangle \in c$, then (82) will map this to $\langle w_F, A'\rangle \in c$ such that for each $a' \in A'$, $a(y) =$ French, and so does (80), because that's the maximal value. Then, all such possibilities $\langle w_F, A'\rangle \in c$ will be removed by the scalar implicature, meaning that we have the inference that it's not the case that everyone speaks only French, which is good.

Crucially, we do not remove all the possibilities involving w_{Fg}. This is because the two sentences will create different possibilities out of $\langle w_{Fg}, A\rangle$. Specifically, (82) will map it to $\langle w_{Fg}, A'\rangle$ where each $a' \in A'$ is such that $a'(y) =$ French, but (80) will map it to $\langle w_{Fg}, A''\rangle$ where for some $a'' \in A''$, $a'(y) =$ French $\oplus$ German, when $a'(x)$

speaks both languages in w_{Fg}. This means, therefore, there will be some possibilities left in the output information state whose world component is w_{Fg}, so the scalar implicature will not entail that not every applicant speaks French.

By the same reasoning on the alternative (74b), we obtain the inference that not every applicant only speaks German. Thus, overall, the reading with all the implicatures factored in will amounts to: Every applicant speaks at least one of French and German, and the following are all false: every applicant only speaks French, every applicant only speaks German, and every applicant speaks both of them. This is compatible with every applicant speaking French as long as only some of them speak German, and also with every applicant speaking German, as long as only some of them speak French. Notice in particular that we do not need pruning or any extra machinery like embedded implicature or innocent inclusion to derive this reading. The crucial ingredient is the discourse referents, which makes the alternatives more informative than usually assumed, thereby making their 'negations' weaker than usually assumed.

There is a remaining question about whether we also want to derive a reading that entails that not every applicant speaks French and that not every applicant speaks German. The two previous analyses, Crnič et al. (2015) and Bar-Lev and Fox (2020), actually derive this stronger reading by default. Under the present account, it cannot be derived as a separate reading, but I am not completely certain if it is actually a separate reading to begin with. That is, the inference that not everyone speaks French and not every speaks German follows from the reading we derived together with an additional assumption that everyone speaks at most one of them. At this point I do not think there is conclusive evidence that the stronger partial plurality inference needs to be represented separately. I will therefore leave this potential issue for my analysis open for now.

Relatedly, Bar-Lev and Fox (2020) observe that only the stronger partial plurality inference is available when the universal quantifier is a universal modal, as in (83).

(83) Daniel is required to speak French or German (at the interview).

The partial plurality inference here amounts to the negations of all the following alternatives.

(84) a. Daniel is required to speak French.
b. Daniel is required to speak German.
c. Daniel is required to speak French and German.

Bar-Lev and Fox (2020) conjecture that this difference between universal DP quantifiers and universal modals arises because the latter do not introduce existential modals as alternatives in this case. Our analysis might be able to account for it by capitalising on the fact that the modal blocks anaphora, as shown in (85).

(85) Daniel is required to speak [French or German]^x at the interview. ??Nathan spoke it$_x$.

If there is no discourse referents, then the present theory predicts the negations of (84) to be the scalar implicatures.

However, one complication here is that modals give rise to a restricted form of anaphora called *modal subordination*, like (86).

(86) Daniel is required to speak [French or German]^x at the interview. The interviewer will ask some simple questions in it$_x$.

Thus, the prediction of the theory needs to be evaluated carefully with respect to modal subordination. I will leave this for future research.

Lastly, the account offered here is much more limited in its empirical scope than its competitors, as I have not offered a general semantics for disjunction, while distributivity inferences are observed with all kinds of disjunction, not just with DP disjunction. Extending the present account to such cases will require considerable change in the system, as the anaphoric mechanism will have to be generalised to all semantic types. This could probably be done, but I will not try to develop such an extension of the theory, to keep the paper at a reasonable length.

7 Conclusion

To the best of my knowledge, the present paper is the first systematic investigation of the idea that scalar implicatures can be computed relative to the anaphoric dimension of meaning, which is represented in terms of discourse referents. As I have remarked multiple times, this is a particularly natural idea given Grice's (1989) intuition that scalar implicatures are drawn from alternatives that would have been more informative (or not less informative), and given that discourse referents carry information. I presented a concrete formal implementation of the idea in dynamic semantics, and discussed its consequences in one empirical domain, the plurality inferences of plural nouns in English and their interactions with quantifiers. I argued that it not only explains the representative data points discussed in the literature, including partial plurality inferences, but also does so in significantly simpler ways, as it requires no additional machinery like higher-order implicatures or embedded implicatures. Furthermore, the theoretical value of the proposal goes beyond this one empirical phenomenon, as it is an idea that is applicable to any phenomena involving discourse referents more generally. I hope to explore further consequences of this idea in other empirical domains in future research.

Appendix A: Different models of discourse referents

Although the term *discourse referent* is more often used in the context of dynamic semantics, discourse referents as abstract objects carrying anaphoric information are amenable to alternative formalisations, and some such ideas can be found in the literature. The reason why I chose dynamic semantics in formalising my analysis in this paper is primarily because of its practical advantage that the interactions between discourse referents and quantifiers, including quantificational subordination, have already been given formal analyses by previous studies, and theoretically speaking, other ways of representing discourse referents could have been used just as well. However, in comparison to dynamic semantics, a lot of theoretical work is yet to be done in such theories of discourse referents, as far as I can see. Here I will mention two alternative ways of modelling discourse referents.

The E-type theory of pronominal anaphora couched in situation semantics (Heim, 1990; Elbourne, 2005) is often contrasted with dynamic semantics. Concretely, Elbourne's (2005) theory would account for the contrast in (1), repeated here, roughly as follows.

(1) a. Paul has a TIN. But he hasn't used it in a while.
 b. Paul is a registered taxpayer. ??But he hasn't used it in a while.

The truth-conditions of the first sentences of these examples are contextually equivalent but they differ in two crucial respects: (i) what kind of situations they mention, and (ii) what kind of noun phrases they make salient. By assumption, a successful use of *it* in the second sentence to refer to Paul's TIN requires a prior mention of a situation that contains a unique TIN, and also the noun *TIN* to be salient in the discourse. Both of these are satisfied in (1a) but not in (1b). In this framework, therefore, we can regard 'discourse referents' as carriers of these two pieces of information.

The debate about the E-type versus dynamic approaches to pronominal anaphora is often phrased in terms of static versus dynamic semantics, but it should be noticed that as far as discourse anaphora like (1) is concerned, there is no way that the E-type system can stay completely static. In other words, the debate is only about whether the meanings of quantificational determiners and certain connectives involved in donkey anaphora need to be inherently dynamic or can stay static, and when it comes to discourse anaphora, any approach requires some dynamic mechanism that passes information from one sentence to another.[31] However, the dynamic aspect of the E-type approach is often not formalised (but see Mandelkern and Rothschild, 2019), and this needs to be worked out before one could build a theory of scalar implicature with discourse referents. As far as I know, there also has been no discussion of quantificational subordination in this framework so far.

Another approach to pronominal anaphora is suggested by Stalnaker (1998: p.11), which is an attempt to stay as conservative as possible with respect to how to model discourse contexts, by accounting for pronominal anaphora solely in terms of sets of possible worlds. The crucial observation is that while the truth-conditions of the first

[31] The debate also tends to ignore anaphora across conjuncts, which needs a certain degree of dynamicisation even within a single sentence. See Mandelkern and Rothschild (2019) for detailed discussion.

sentences of (1) are indeed contextually equivalent, their assertions do not result in the exact same common ground. That is, after the utterance of the first sentences, the common ground will be slightly different between the two cases, because in the first case, it will be commonly known that the speaker has uttered the first sentence of (1a), while in the second case, it will be commonly known that the speaker has uttered the first sentence of (1b). Stalnaker conjectures that this difference makes a difference with respect to pronominal anaphora. Under this view, therefore, we could think of discourse referents abstractly in terms of such differences in the set of possible common grounds. However, it has not been fully specified exactly how such information is used in resolving pronominal anaphora, let alone how quantificational subordination is to be analysed.

Appendix B: The homogeneity approach

(2017) puts forward a homogeneity-based analysis of plurality inferences. Homogeneity inferences are typically observed with expressions like definite plurals as in (A-1).

(A-1) Aygül read the books.

A proper characterisation of the truth-conditions of this sentence seems to require reference to (at least) three truth-values: It is true when Aygül read all the relevant books, and is false when she didn't read any of these books, but when neither of these is the case, one normally hesitates to say that the sentence is clearly true or clearly false. This is perhaps a gradient phenomenon, but for the sake of exposition, let us follow Križ and treat such judgments with a third truth-value, which I denote by #.[32] Thus, under this analysis, the truth-conditions of (A-1) can be represented as follows. We ignore the presuppositions in order to simplify the discussion.

(A-2) $$\text{(A-1) denotes} \begin{cases} 1 & \text{if Aygül read all of the books} \\ 0 & \text{if Aygül read none of the books} \\ \# & \text{otherwise} \end{cases}$$

Križ analysis of plurality inferences of plural noun phrases is a reductionist one in that he treats it as a homogeneity inference. Specifically, just like definite plurals give rise to trivalent truth-conditions, he proposes that indefinite plurals also give rise to trivalent truth-conditions, as illustrated in (A-3).

(A-3) $$\text{‘Aygül read books’ denotes} \begin{cases} 1 & \text{if Aygül read multiple books} \\ 0 & \text{if Aygül read no book} \\ \# & \text{otherwise} \end{cases}$$

[32] I will also ignore non-trivalent approaches like Bar-Lev (2018), as it is not my purpose here to argue for a particular analysis of homogeneity, but to argue that plurality inferences are not to be understood in terms of homogeneity.

This analysis looks plausible given parallel behavior of definite and indefinite plurals in the scope of certain operators. For instance, for both of them, negation gives rise to meaning stronger than what would be expected under the bivalent analysis. This can be captured in the current trivalent setting by assuming that negation preserves trivalency as follows.

(A-4) $$\text{`not } \phi\text{' denotes } \begin{cases} 1 & \text{if `}\phi\text{' denotes } 0 \\ 0 & \text{if `}\phi\text{' denotes } 1 \\ \# & \text{otherwise} \end{cases}$$

The homogeneity approach to plurality inferences has certain attractive features, but issues arise when quantification enters the picture. Based on his earlier work on definite plurals, Križ (2017) deals with quantifiers based on their bivalent meaning Q with the help of auxiliary operators that turn tivalent meaning bivalent, which I call $\mathcal{A}$ and $\mathcal{B}$ here. For any function P of type $\langle e, t\rangle$:

(A-5) a. $$\mathcal{A}(P) = \lambda x. \begin{cases} 1 & \text{if } P(x) = 1 \\ 0 & \text{otherwise} \end{cases}$$

b. $$\mathcal{B}(P) = \lambda x. \begin{cases} 0 & \text{if } P(x) = 0 \\ 1 & \text{otherwise} \end{cases}$$

Then for any bivalent quantifier Q of type $\langle et, t\rangle$, the trivalency of P projects as follows.

(A-6) $$Q(P) = \begin{cases} 1 & \text{if } Q(\mathcal{A}(P)) = Q(\mathcal{B}(P)) = 1 \\ 0 & \text{if } Q(\mathcal{A}(P)) = Q(\mathcal{B}(P)) = 0 \\ \# & \text{otherwise} \end{cases}$$

Thus, Križ (2017) does not give quantifiers trivalent meaning but provide a general projection rule based on their bivalent meaning.

Let us apply the above recipe to *exactly one*. Consider the following sentences.

(A-7) a. Exactly one student read the books.
b. Exactly one student read books.

These sentences are predicted to have the following denotations:

(A-8) a. $$\text{(A-7a) denotes } \begin{cases} 1 & \text{if one student read all the books,} \\ & \quad \text{and every other student read none of the books} \\ 0 & \text{if no student read any of the books,} \\ & \quad \text{or more than one student read all the books,} \\ & \quad \text{or more than one student read at least one of the books} \\ & \qquad \text{but no student read all the books} \\ \# & \text{otherwise} \end{cases}$$

$$\text{b. (A-7b) denotes} \begin{cases} 1 & \text{if one student read multiple books,} \\ & \quad \text{and every other student read no books} \\ 0 & \text{if no student read any book,} \\ & \quad \text{or more than one student read multiple books,} \\ & \quad \text{or more than one student read exactly one book} \\ & \qquad \text{and no student read multiple books} \\ \# & \text{otherwise} \end{cases}$$

The falsity conditions of these sentences are quite complicated, and I am not completely sure if the third disjuncts in the above reprsentation are empirically justifiable, but I refrain from assessing the analysis based on this point. See Križ (2017, §5.2) for relevant discussion.

It seems to me that a more grave issue for this approach to plurality inferences comes from the observation that the behavior of definite and indefinite plurals diverge in the scope of *every*. For example, consider (A-9).

(A-9) a. Every student read the books.
b. Every student read books.

As discussed in the main text, (A-9a) seems to have a reading with a partial plurality inference that at least some of the students read multiple books, in addition to the fully plural reading that every student read multiple books. On the other hand, (A-9a) does not seem to be ambiguous in the same way; rather, it seems to entail that every student read the same set of multiple books, and they read all of them. This contrast is not accounted for by Križ's (2017) analysis.

To be more precise, let us examine his predictions. The denotations of these sentences will be as in (A-10).

$$\text{(A-10) a. (A-9a) denotes} \begin{cases} 1 & \text{if every student read all the books} \\ 0 & \text{if at least one student read none of the books} \\ \# & \text{otherwise} \end{cases}$$

$$\text{b. (A-9b) denotes} \begin{cases} 1 & \text{if every student read multiple books} \\ 0 & \text{if at least one student read no book} \\ \# & \text{otherwise} \end{cases}$$

Above all, notice that the predicted meaning for (A-9b) is a strong one, entailing that every student read multiple books, rather than the weak reading where at least some of them read multiple books. Thus, the predicted default reading is a strong, fully plural one, unlike for my theory.

Križ (2017) is aware of this, and proposes to derive the weak reading by referring to the context sensitivity of homogeneity. This is a well motivated assumption, particularly for the homogeneity of definite plurals (see Križ, 2016; Križ and Spector, 2021 and references), but crucially Križ predicts that it applies to the plurality inferences of plural indefinites as well. Specifically, he follows his earlier work, Križ (2016), and assumes that given a sentence with a non-trivially trivalent denotation, certain possible worlds in which it denotes # can may be regarded as if they make the sentence true, if the distinction between these two types of worlds is not pragmatically important. Then he remarks as follows of the sentence *Every girl saw zebras*, which has a denotation similar to (A-10b):

> It is quite possible, for example, that in some context, we may not care about how many girls, exactly, saw multiple zebras as long as some of them did, in which case the overall interpretation would be that *every girl saw one or more zebras and some saw more than one*. (Križ, 2017: p. 28).

I find this explanation unsatisfactory because in such a context, the definite version should also give rise to a 'partial reading' too, which will only entail that at least some of the students read all of the relevant books. However, this does not seem to be true. In particular, if the partial plurality reading of (A-9b) is made possible by a contextual assumption that it does not matter how many students read multiple books as long as some did and every student read at least one book, then in the same context, one should be able to truthfully utter (A-9a), even if some students read only one of the books, contrary to fact. To further reinforce this point, consider the following scenario. Students are required to read a book in order to pass, and teachers are rewarded if there are some who have read multiple books. One of the teachers says to another:

(A-11) Every student in your class read books, so you will be rewarded. Every student in my class read these books, so I will be rewarded too.

It is reasonable to assume that the relevant contextual factors are constant between the two sentences, but contrary to Križ's (2017) prediction, the partial reading is only available for the first sentence with an indefinite plural, and not for the second sentence with a definite plural. It's possible that definite plurals may also give rise to partial readings in certain contexts, but it appears to me that the licensing conditions for them are not the same as for the partial plurality inferences observed with indefinites, which are much more easily accessible.

For this reason, I do not think Križ's (2017) homogeneity theory gives a full account of partial plurality inferences. It is certainly a possibility that a different version of the homogeneity theory, perhaps with a different analysis of quantifiers, can explain the above contrast between indefinite and definite plurals, but until such a theory has been proposed, the scalar implicature approach remains to be the only approach that can satisfactorily deal with partial plurality inferences.

References

Anand, P., Andrews, C., Farkas, D., & Wagers, M. (2011). The exclusive interpretation of plural nominals in quantificational environments. In N. Ashton, A. Chereches, & D. Lutz (Eds.), *Proceedings of SALT 21* (pp. 176–196). https://doi.org/10.3765/salt.v21i0.2617.

Anvari, A. (2019). *Aspects of contextual enrichment*. Ecole Normale Supérieure dissertation.

Bale, A., & Barner, D. (2009). The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics, 26*(3), 217–252. https://doi.org/10.1093/jos/ffp003.

Bale, A., Gagnon, M., & Khanjian, H. (2011). On the relationship between morphological and semantic markedness. *Morphology, 21*(2), 197–221. https://doi.org/10.1007/s11525-010-9158-1.

Bar-Lev, M. (2018). *Free choice, homogeneity and innocent inclusion*. The Hebrew University of Jerusalem dissertation.

Bar-Lev, M., & Fox, D. (2020). Free choice, simplification, and Innocent Inclusion. *Natural Language Semantics, 28*, 175–223. https://doi.org/10.1007/s11050-020-09162-y.

Barner, D., & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition, 97*(1), 41–66. https://doi.org/10.1016/j.cognition.2004.06.009.

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy, 4*(2), 159–219.

Beaver, D. (2001). *Presupposition and assertion in dynamic semantics*. Stanford: CSLI Publications.

Beaver, D., & Krahmer, E. (2001). A partial account of presupposition projection. *Journal of Logic, Language, and Information, 10*(2), 147–182. https://doi.org/10.1023/A:1008371413822.

Bergen, L., Levy, R., & Goodman, N. (2016). Pragmatic reasoning through semantic inference. *Semantics & Pragmatics, 9*(20), 1–83. https://doi.org/10.3765/sp.9.20.

Brasoveanu, A. (2007). *Structured nominal and modal reference*. Rutgers University dissertation.

Brasoveanu, A. (2008). Donkey pluralities: Plural information states versus non-atomic individuals. *Linguistics and Philosophy, 31*(2), 129–209. https://doi.org/10.1007/s10988-008-9035-0.

Breheny, R., Klinedinst, N., Romoli, J., & Sudo, Y. (2018). The symmetry problem: Current theories and prospects. *Natural Language Semantics, 26*(2), 85–110. https://doi.org/10.1007/s11050-017-9141-z.

Buccola, B., & Spector, B. (2016). Modified numerals and maximality. *Linguistics and Philosophy, 39*, 151–199. https://doi.org/10.1007/s10988-0169187-2.

Bylinina, L., & Nouwen, R. (2018). On “zero" and semantic plurality. *Glossa, 3*(1), 98. https://doi.org/10.5334/gjgl.441.

Champollion, L., Bumford, D., & Henderson, R. (2019). Donkeys under discussion. *Semantics and Pragmatics*. https://doi.org/10.3765/sp.12.1.

Charlow, S. (2014). *On the semantics of exceptional scope*. New York University dissertation.

Chemla, E. (2008). An epistemic step for anti-presuppositions. *Journal of Semantics, 25*(2), 141–173. https://doi.org/10.1093/jos/ffm017.

Chemla, E. (2009). *Similarity: Towards a unified account of scalar implicatures, free choice permission and presupposition projection*. LSCP and MIT: Ms.

Chemla, E., & Spector, B. (2011). Experimental evidence for embedded scalar implicatures. *Journal of Semantics, 28*(3), 359–400. https://doi.org/10.1093/jos/ffq023.

Chierchia, G. (1995). *Dynamics of meaning: Anaphora, presupposition, and the theory of grammar*. Chicago: University of Chicago Press.

Chierchia, G. (1998). Plurality of mass nouns and the notion of ‘semantic parameter’. In S. Rothstein (Ed.), *Events and grammar* (pp. 53–103). Dordrecht: Kluwer.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond: The cartography of syntactic structures* (pp. 39–103). Oxford: Oxford University Press.

Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese, 174*(1), 99–149. https://doi.org/10.1007/s11229-009-9686-6.

Chierchia, G., Fox, D., & Spector, B. (2012). Scalar implicature as a grammatical phenomenon. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 3, pp. 2297–2331). Berlin: de Gruyter.

Clifton, C., & Dube, C. (2010). Embedded implicatures observed: A comment on Geurts and Pouscoulous (2009). *Semantics & Pragmatics, 3*(7), 1–13. https://doi.org/10.3765/sp.3.7.

Crnič, L., Chemla, E., & Fox, D. (2015). Scalar implicatures of embedded disjunction. *Natural Language Semantics, 23*(4), 271–305. https://doi.org/10.1007/s11050-015-9116-x.

Cummins, C. (2014). Typicality made familiar: A commentary on Geurts and van Tiel (2013). *Semantics & Pragmatics, 7*(8), 1–15. https://doi.org/10.3765/sp.7.8.

Dotlačil, J. (2013). Reciprocals distribute over information states. *Journal of Semantics, 30*(4), 423–477. https://doi.org/10.1093/jos/ffs016.

Egré, P., & Cozic, M. (2016). Conditionals. In M. Aloni & P. Dekker (Eds.), *The Cambridge handbook of formal semantics* (pp. 490–524). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139236157.018.

Elbourne, P. (2005). *Situations and individuals*. Cambridge, MA: MIT Press.

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

Farkas, D. F., & de Swart, H. E. (2010). The semantics and pragmatics of plurals. *Semantics & Pragmatics, 3*(6), 1–54. https://doi.org/10.3765/sp.3.6.

Fox, D. (2007). Free choice and the theory of scalar implicatures. In U. Sauerland & P. Stateva (Eds.), *Presupposition and implicature in compositional semantics* (pp. 71–112). Basingstoke: Palgrave Macmillan.

Fox, D. (2012). Presupposition projection from quantificational sentences: Trivalence, local accommodation, and presupposition strengthening. In I. Caponigro & C. Cecchetto (Eds.), *From grammar to meaning: The spontaneous logicality of language* (pp. 201–232). Cambridge: Cambridge University Press.

Fox, D., and Hackl, M. (2006). The universal density of measurement. *Linguistics and Philosophy, 29*(5), 537–586. https://doi.org/10.1007/s10988-006-9004-4.

Fox, D., & Katzir, R. (2011). On the characterization of alternatives. *Natural Language Semantics, 19*(1), 87–107. https://doi.org/10.1007/s11050-010-9065-3.

Fox, D., and Katzir, R. (2021). Notes on iterated rationality models of scalar implicatures. *Journal of Semantics, 38*(4), 571–600. https://doi.org/10.1093/jos/ffab015.

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

Franke, M., & Bergen, L. (2020). Theory-driven statistical modeling for semantics and pragmatics: A case study on grammatically generated implicature readings. *Language, 96*(2), 77–96. https://doi.org/10.1353/lan.2020.0034.

Franke, M., Schlotterbeck, F., & Augurzky, P. (2017). Embedded scalars, preferred readings and prosody: An experimental revisit. *Journal of Semantics, 34*(1), 153–199. https://doi.org/10.1093/jos/ffw007.

Gajewski, J., & Sharvit, Y. (2012). In defense of the grammatical approach to local implicatures. *Natural Language Semantics, 20*(1), 31–57. https://doi.org/10.1007/s11050-011-9074-x.

George, B. (2008). A new predictive theory of presupposition projection. In T. Friedman & S. Ito (Eds.), *Proceedings of SALT 18* (pp. 358–375). https://doi.org/10.3765/salt.v18i0.2472.

Geurts, B. (2008). Implicture as a discourse phenomenon. In L. McNally & E. Puig-Waldmüller (Eds.), *Proceedings of Sinn und Bedeutung* (pp. 261–275). https://doi.org/10.18148/sub/2007.v11i0.644.

Geurts, B. (2009). Scalar implicature and local pragmatics. *Mind & Language, 24*(1), 51–79. https://doi.org/10.1111/j.1468-0017.2008.01353.x.

Geurts, B., & Pouscoulous, N. (2009). Embedded implicatures? *Semantics & Pragmatics, 2*(4), 1–34. https://doi.org/10.3765/sp.2.4.

Geurts, B., & van Tiel, B. (2013). *Embedded scalars. Semantics & Pragmatics, 6*(9), 1–37. https://doi.org/10.3765/sp.6.9.

Grice, P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Groenendijk, J., & Stokhof, M. (1990). Dynamic Montague Grammar. In L. Karttunen & L. Polos (Eds.), *Proceedings of the 2nd Symposium on Logic and Language* (pp. 3–48). Budapest: Eotvos Lorand University Press.

Grimm, S. (2013). Plurality is distinct from number-neutrality. In Y. Fainleib, N. LaCara & Y. Park (Eds.), *Proceedings of NELS 41* (pp. 247–258). Amherst: GLSA.

Groenendijk, J., & Stokhof, M. (1991). Dynamic predicate logic. *Linguistics and Philosophy, 14*(1), 39–100. https://doi.org/10.1007/BF00628304.

Gutzmann, D. (2015). *Use-conditional meaning: Studies in multidimensional semantics*. Oxford: Oxford University Press.

Gutzmann, D., & McCready, E. (2016). Quantification with pejoratives. In R. Finkbeiner, J. Meibauer, & H. Wiese (Eds.), *Pejoration* (pp. 75–102). Amsterdam: John Benjamins.

Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. University of Massachusetts, Amherst dissertation.

Heim, I. (1983). On the projection problem for presuppositions. In M. Barlow, D. P. Flickinger & M. Wescoat (Eds.), *Proceedings of WCCFL 2* (pp. 114–125). Stanford: CSLI Publications.

Heim, I. (1990). E-type pronouns and donkey anaphora. *Linguistics and Philosophy, 13*(2), 137–177. https://doi.org/10.1007/BF00630732.

Heim, I. (1991). Artikel und Definitheit. In A. von Stechow & D. Wunderlich (Eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung,* (pp. 487–535). Berlin: de Gruyter.

Heim, I. (2011). Definiteness and indefiniteness. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 2, pp. 996–1025). Berlin: de Gruyter.

Horn, L. (1989). *A natural history of negation*. Chicago: Chicago University Press.

Ivlieva, N. (2013). *Scalar implicatures and the grammar of plurality and disjunction*. Massachusetts Institute of Technology dissertation.

Ivlieva, N. (2014). Multiplicity and non-monotonic environments. In L. Crnič& U. Sauerland (Eds.), *The art and craft of semantics: A festschrift for Irene Heim* (Vol. 1, pp. 245–251). Cambridge, MA: MITWPL. https://semanticsarchive.net/Archive/jZiNmM4N/.

Ivlieva, N. (2020). Dependent plurality and the theory of scalar implicatures: Remarks on Zweig 2009. *Journal of Semantics, 37*(3), 425–454. https://doi.org/10.1093/jos/ffaa004.

Kadmon, N. (1987). *On unique and non-unique reference and asymmetric quantification*. University of Massachusetts, Amherst dissertation.

Kamp, H. (1981). A theory of truth and semantic representation. In J. A. G. Groenendijk, T. M. V. Janssen, & M. J. B. Stokhof (Eds.), *Formal methods in the study of language* (pp. 277–322). Amsterdam: Mathematical Center.

Kamp, H., and Reyle, U. (1993). *From discourse to logic: An introduction to model theoretic semantics in natural language, formal logic and discourse representation theory*. Dordrecht: Kluwer.

Kanazawa, M. (1993). *Dynamic generalized quantifiers and monotonicity*. Report LP-199302. Amsterdam: ILLC.

Kanazawa, M. (1994). Weak vs. strong readings of donkey sentences and mono-tonicity inferences in a dynamic setting. *Linguistics and Philosophy, 17*(2), 109–158. https://doi.org/10.1007/BF00984775.

Karttunen, L. (1976). Discourse referents. In J. D. McCawley (Ed.), *Syntax and semantics 7: Notes from the linguistic underground* (pp. 363–385). New York: Academic Press.

Karttunen, L., & Peters, S. (1979). Conventional implicature. In O. Choon-Kyu & D. Dinneen (Eds.), *Syntax and semantics 11: Presupposition* (pp. 1–56). New York: Academic Press.

Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy, 30*(6), 669–690. https://doi.org/10.1007/s10988-008-9029-y.

Kaufmann, M., & Kaufmann, S. (2015). Conditionals and modality. In S. Lappin & C. Fox (Eds.), *The handbook of contemporary semantic theory* (pp. 237–270). New York: Wiley.

Krahmer, E., & Muskens, R. (1995). Negation and disjunction in discourse representation theory. *Journal of Semantics, 12*(4), 357–376. https://doi.org/10.1093/jos/12.4.357.

Križ, M. (2016). Homogeneity, non-maximality, and 'all'. *Journal of Semantics, 33*(3), 493–539. https://doi.org/10.1093/jos/ffv006.

Križ, M. (2017). Bare plurals, multiciplicity, and homogeneity. Ms., Institut Jean Nicod.

Križ, M., & Spector, B. (2021). Interpreting plural predication: Homogeneity and non-maximality. *Linguistics and Philosophy, 44*, 1131–1178. https://doi.org/10.1007/s10988-020-09311-w.

Landman, F. (2000). *Events and plurality*. Dordrecht: Kluwer.

Landman, F. (2011). Count nouns-mass nouns, neat nouns-mess nouns. *Baltic International Yearbook of Cognition, Logic and Communication, 6*, 1–67. https://doi.org/10.4148/biyclc.v6i0.1579.

Lima, S. (2018). New perspectives on the count-mass distinction: Understudied languages and psycholinguistics. *Language and Linguistics Compass, 12*(11), 1–20. https://doi.org/10.1111/lnc3.12303.

Link, G. (1983). The logical analysis of plurals and mass terms: A lattice theoretical approach. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use, and the interpretation of language* (pp. 302–323). Berlin: Mouton de Gruyter.

Magri, G. (2009a). *A theory of individual-level predicates based on blind mandatory implicatures. Constraint promotion for optimality theory*. Massachusetts Institut of Technology dissertation.

Magri, G. (2009). A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics, 17*(3), 245–297. https://doi.org/10.1007/s11050-009-9042-x.

Mandelkern, M., & Rothschild, D. (2019). Definiteness projection. *Natural Language Semantics, 28*, 77–109. https://doi.org/10.1007/s11050-019-091592.

Martí, L. (2020). Inclusive plurals and the theory of number. *Linguistic Inquiry, 51*(1), 37–74. https://doi.org/10.1162/ling_a_00330.

Marty, P., Romoli, J., Sudo, Y., & Breheny, R. (2022). Negative free choice. *Semantics & Pragmatics*. https://doi.org/10.3765/sp.14.13.

Marty, P. P. (2017). *Implicatures in the DP domain*. Massachusetts Institute of Technology dissertation.

Mayr, C. (2015). Plural definite NPs presuppose multiplicity via embedded exhaustification. In S. D'Antonio, M. Moroney & C. Rose Little (Eds.), *Proceedings of SALT 25* (pp. 204–224). https://doi.org/10.3765/salt.v25i0.3059.

McCready, E. (2010). Varieties of conventional implicatures. *Semantics & Pragmatics, 3*(8), 1-57. https://doi.org/10.3765/sp.3.8.

Meyer, M.-C., & Feiman, R. (2021). Priming reveals similarities and differences between three purported cases of implicature: Some, number and free choice disjunctions. *Journal of Memory and Language*. https://doi.org/10.1016/j.jml.2020.104206.

Muskens, R. (1996). Combining Montague semantics and discourse representation. *Linguistics and Philosophy, 19*(2), 143–186.

Nouwen, R. (2003). *Plural pronominal anaphora in context: Dynamic apsects of quantification*. Utrecht Institute of Linguistics OTS dissertation.

Nouwen, R. (2007). On dependent pronouns and dynamic semantics. *Journal of Philosophical Logic, 36*(2), 123–154. https://doi.org/10.1007/s10992-0069029-8.

Pearson, H., Khan, M., & Snedeker, J. (2010). Even more evidence for the emptiness of plurality: An experimental investigation of plural interpretation as a species of implicature. In N. Li & D. Lutz (Eds.), *Proceedings of SALT 20* (pp. 489–508). https://doi.org/10.3765/salt.v20i0.2554.

Percus, O. (2006). Anti-presuppositions. In A. Ueyama (Ed.), *Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as an empirical science* (pp. 52–73). Washington, DC: Japan Society for the Promotion of Science.

Peters, S., & Westerståhl, D. (2006). *Quantifiers in language and logic*. Oxford: Oxford University Press.

Postal, P. (1966). On so-called 'pronouns' in English. In F. Dinneen (Ed.), *Report on the seventeenth annual round table meeting on linguistics and language studies* (pp. 177–206). Washington, DC: Georgetown University Press.

Potts, C. (2005). *The logic of conventional implicatures*. Oxford: Oxford University Press.

Potts, C. (2007). The expressive dimension. *Theoretical Linguistics, 33*(2), 165–198.

Potts, C., Lassitert, D., Levy, R., & Frank, M. C. (2016). Embedded implicatures as pragmatic inferences under compositional lexical uncertainty. *Journal of Semantics, 33*(4), 755–802. https://doi.org/10.1093/jos/ffv012.

Romoli, J. (2012). *Soft but strong: Neg-raising, soft triggers, and exhaustification*. Harvard University dissertation.

Roth, M. (1987). Noun phrase interpretation in Montague grammar, file change semantics, and situation semantics. In P. Gärdenfors (Ed.), *Generalized quantifiers: Linguistic and logical approaches* (pp. 237–268). Dordrecht: Reidel.

Roth, M., & Partee, B. (1982). Conjunction type ambiguity, and wide scope *or*. In D. P. F. Marlys Macken & N. Wiegand (Eds.), *Proceedings of WCCFL 1* (pp. 353–362). Stanford, CA: CSLI Publications.

Rothschild, D. (2017). A trivalent approach to anaphora and presupposition. In A. Cremers, T. van Gessel & F. Roelofsen (Eds.), *Proceedings of the Amsterdam Colloquium 2017* (pp. 1–13). Amsterdam: ILLC.

Rothschild, D., & Yalcin, S. (2016). Three notions of dynamicness in language. *Linguistics and Philosophy, 39*(4), 333–355. https://doi.org/10.1007/s10988-016-9188-1.

Rothstein, S. (2017). *Semantics for counting and measuring*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9780511734830.

Sauerland, U. (2003). A new semantics for number. In R. B. Young & Y. Zhou (Eds.), *Proceedings of SALT 13* (pp. 258–275). https://doi.org/10.3765/salt.v13i0.2898.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy, 27*(3), 367–391. https://doi.org/10.1023/B:LING.0000023378.71748.db.

Sauerland, U. (2008). Implicated presuppositions. In A. Steube (Ed.), *The discourse potential of underspecified structures* (pp. 581–600). Berlin: Mouton de Gruyter.

Sauerland, U., Anderssen, J., & Yatsushiro, K. (2005). The plural is semantically unmarked. In S. Kepser & M. Reis (Eds.), *Linguistic evidence* (pp. 409–430). Berlin: Mouton de Gruyter.

Schlenker, P. (2006). Scopal independence: A note on branching and wide scope readings of indefinites and disjunctions. *Journal of Semantics, 23*(3), 281–314. https://doi.org/10.1093/jos/ffl005.

Schlenker, P. (2012). Maximize presupposition and Gricean reasoning. *Natural Language Semantics, 20*(4), 391–429. https://doi.org/10.1007/s11050-0129085-2.

Schwarzschild, R. (1996). *Pluralities*. Dordrecht: Kluwer.

Singh, R. (2019). Context, content, and the occasional costs of implicature computation. *Frontiers in Psychology, 10*, 2214. https://doi.org/10.3389/fpsyg.2019.02214.

Spector, B. (2006). *Aspects de la pragmatique des operateurs logiques*. Université Paris 7, Denis Diderot dissertation.

Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In U. Sauerland & P. Stateva (Eds.), *Presuppositions and implicatures in compositional semantics* (pp. 243–281). New York: Palgrave-Macmillan.

Spector, B. (2016). *Comparing exhaustivity operator. Semantics & Pragmatics, 9*(11), 1–33. https://doi.org/10.3765/sp.9.11.

Spector, B., & Sudo, Y. (2017). Presupposed ignorance and exhaustification: How scalar implicatures and presuppositions interact. *Linguistics and Philosophy, 40*(5), 473–517. https://doi.org/10.1007/s10988017-9208-9.

Stalnaker, R. (1998). On the representation of context. *Journal of Logic, Language and Information, 7*, 3–19.

Stateva, P., Andreetta, S., & Stepanov, A. (2016). On the nature of the plurality inference: Ladybugs for Anne. In L. Dupuy, A. Grabizna, N. Foudon, & P. Saint-Gernier (Eds.), *Papers dedicated to Anne Reboul.* Lyon: CNRS.

Stone, M. (1992). 'Or' and anaphora. In C. Barker & D. Dowty. (Eds.), *Proceedings of SALT 2* (pp. 367–385). Columbus, OH: The Ohio State University working papers in linguistics.

Sudo, Y. (2012). *On the semantics of phi features on pronouns*. Massachusetts Institute of Technology dissertation.

Sudo, Y. (2014). Presupposition projection in quantified sentences and cross-dimensional anaphora. Ms., University College London.

Sudo, Y. (2016). The existential problem of scalar implicatures and anaphora across alternatives. In C. Piñon (Ed.), *Empirical issues in formal syntax and semantics* (Vol. 11, pp. 225–244). http://www.cssp.cnrs.fr/eiss11/index_en.html.

Tieu, L., Bill, C., Romoli, J., & Crain, S. (2020). Testing theories of plural meanings: Insights from acquisition. *Cognition, 205*. https://doi.org/10.1016/j.cognition.2020.104307.

Tieu, L. & Romoli, J. (2019). Plurality. In C. Cummins & N. Katsos (Eds.), *The Oxford handbook of experimental semantics and pragmatics* (pp. 208–227). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198791768.013.22.

van den Berg, M. (1996). *Some aspects of the internal structure of discourse: The dynamics of nominal anaphora*. Universiteit van Amsterdam dissertation.

van Eijck. (2001). Incremental dynamics. *Journal of Logic, Language, and Information, 10*(3), 319–351. https://doi.org/10.1023/A:1011251627260.

von Fintel, K. (2011). Conditionals. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.), *Semantics* (Vol. 2, pp. 1515–1438). Berlin: de Gruyter.

von Fintel, K. (2012). Subjunctive conditionals. In G. Russell & D. G. Fara (Eds.), *The Routledge comparion to philosophy of language* (pp. 466–477). New York: Routledge.

van Tiel, B. (2014). Embedded scalars and typicality. *Journal of Semantics, 31*(2), 147–177. https://doi.org/10.1093/jos/fft002.

van Tiel, B., Noveck, I. A., & Kissine, M. (2018). Reasoning with 'some'. *Journal of Semantics, 35*(4), 757–797. https://doi.org/10.1093/jos/ffy012.

van Tiel, B., & Pankratz, E. (2021). Adjectival polarity and the processing of scalar inferences. *Glossa, 6*(1), 32. https://doi.org/10.5334/gjgl.1457.

van Tiel, B., Pankratz, E., & Sun, C. (2019). Scales and scalarity: Processing scalar inferences. *Journal of Memory and Language, 105*, 93–107. https://doi.org/10.1016/j.jml.2018.12.002.

van Tiel, B., van Miltenburg, E., Zevakhina, N., & Geurts, B. (2016). Scalar diversity. *Journal of Semantics, 33*(1), 137–175. https://doi.org/10.1093/jos/ffu017.

Winter, Y. (2001). *Flexibility principles in Boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. Cambridge, MA: MIT Press.

Zweig, E. (2009). Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy, 32*(4), 353–407. https://doi.org/10.1007/s10988009-9064-3.