On fatal competition and the nature of distributive inferences*

Moshe E. Bar-Lev and Danny Fox

TAU and MIT

January 10, 2023

Abstract

Denić (2021) observes that the availability of distributive inferences — for sentences with disjunction embedded in the scope of a universal quantifier — depends on the size of the domain quantified over as it relates to the number of disjuncts. Based on her observations, she argues that probabilistic considerations play a role in the computation of implicatures. In this paper we explore a different possibility. We argue for a modification of Denić's generalization, and provide an explanation that is based on intricate logical computations but is blind to probabilities. The explanation is based on the observation that when the domain size is no larger than the number of disjuncts, universal and existential alternatives are equivalent if distributive inferences are obtained. We argue that under such conditions a general ban on 'fatal competition' (Magri 2009a,b; Spector 2014) is activated thereby predicting distributive inferences to be unavailable.

1 Introduction

An utterance of the sentence in (1) normally leads to the inference that some student read War and Peace and that some (other) student read Brothers Karamazov. These two inferences, often referred to as Distributive Inferences (DIs), are normally assumed to be derived as Scalar Implicatures (SIs).

(1) Every student read War and Peace or Brothers Karamazov.

Distributive Inferences (DIs):

a. $\rightsquigarrow$ some student read War and Peace.

*For helpful discussions and comments we thank Amir Anvari, Luka Crnič, Milica Denić, Roni Katzir, Benjamin Spector, and the audience at the LINGUAE Seminar in Paris. We also wish to acknowledge Keny Chatain's Exh package (https://github.com/KenyC/Exh). His code was very helpful to us in understanding the consequences of various conjectures we made as we developed our proposal. Moshe Bar-Lev was supported by a postdoctoral fellowship from the Cukier-Goldstein-Goren Center for Mind and Language.

b. $\rightsquigarrow$ some student read Brothers Karamazov.

Denić (2021) argues for a very interesting claim about the distribution of DIs, one that connects the number of disjuncts to the cardinality of the quantificational domain. Specifically, she argues for the generalization in (2).

(2) **The Denić generalization**:
In sentences of the form $\forall x \in D(P_1 x \vee \ldots \vee P_n x)$, preference for deriving distributive inferences gradually decreases as $|D| - n$ decreases.

As evidence for this generalization, she contrasts the sentences in (3) and claims that the availability of DIs indeed decreases as the difference between $|D|$ and n decreases: The DIs are most accessible for (3a), least accessible for (3c), and somewhere in the middle for (3b).

(3) a. All of Mary's twenty friends are French or Spanish. $|D| - n = 18$
b. All of Mary's four friends are French or Spanish. $|D| - n = 2$
c. Both of Mary's friends are French or Spanish. $|D| - n = 0$

She, then, provides an explanation for her generalization in (2) based on the assumption that probabilities enter into the computation of SIs, specifically that they affect the contextual choice of alternatives, among those formally determined by grammar (that it enters into, so called, pruning). The two goals of this paper are:

1. to argue that there is a categorical constraint that allows DIs if and only if $|D| > n$, and
2. to provide a modular account for the constraint — one which is based on a principle underlying what we call fatal competition, a principle that examines intricate logical properties of linguistic expression but not their probabilities.

If these goals were met, there would be a modular non-probabilistic account for (3c). The remaining question is whether there would still be an argument (albeit weaker) in favor of Denić's generalization based on the contrast between (3a) and (3b), thereby possibly still providing evidence for her probabilistic account. While this is not the main goal of our paper, we will suggest that there is indeed a probabilistic account for the contrast, but one which does not pertain to the computation of SIs, a computation that remains blind to probabilities. Instead, it pertains to disambiguation, something which is known to interact with plausibility considerations in complex ways. Specifically, within the grammatical theory of SIs, the presence or absence of an SI is determined by the parse chosen for the relevant linguistic stimulus, a choice that is bound to be affected by a variety of extra-linguistic considerations, some of which pertain to the relative likelihood of various states of affairs.

The structure of the paper is as follows. In §2 we argue for the alternative categorical generalization we mentioned above. In §3 we lay out our assumptions as to how DIs are derived (following Bar-Lev and Fox 2016, 2020), and in §4 we point out that this derivation is completely parallel to the derivation of free choice inferences (FCIs; Fox 2007). Then, in §5, we argue that with these derivations of DIs and FCIs it is predicted that when $n \geq |D|$ sentences of the form $\forall x \in D(P_1 x \vee \ldots \vee P_n x)$ fatally compete with their existential alternative. In §6 we propose that the unacceptability of cases where there is fatal competition follows from a constraint requiring the alternatives of a sentence to identify a question, potentially with the aid of exhaustification (building on Katzir and Singh 2015; Fox 2018; Anvari 2022). §7 discusses some further predictions of our proposal, and §8 concludes while providing a sketch of an account of the intermediate status of (3b); specifically, we explain why this intermediate status, which is not predicted by our proposal, is nevertheless compatible with it, and in fact derivable with the aid of straightforward assumptions about the probabilistic considerations that could enter into disambiguation.

2 The Categorical Denić Generalization

Our goal in this section is to demonstrate that there is a categorical effect which is not captured by Denić's generalization in (2). Simple evidence for this comes from the minimal pair in (4).

(4) a. (I am sad that I can't see my three children on a regular basis.) I live in Israel and they each live in Europe or the United States.

b. #(I am sad that I can't see my two children on a regular basis.) I live in Israel and they both live in Europe or the United States.

(4a) is entirely acceptable with a DI, despite the fact that $|D| - n$ is as small as it can be with $|D| > n$. (4b), by contrast, doesn't allow for a DI to arise. Instead it obligatorily leads to the inference that the speaker is not (entirely) opinionated about where her children live.[1] The sentence sounds odd because this is not particularly plausible (but simple variants sound better if one imagines that the speaker intends to present herself as a detached parent, e.g., *I am lucky I don't need to see my children that often, they live somewhere far away. Both live either in Europe or the United States*).

This categorical contrast supports the claim that DIs require the size of the quantification domain to be larger than the number of disjuncts. We thus state a categorical version of Denic's generalization in (5) (Denić entertains a similar non-gradable generalization, but takes the intermediate

1 The source of such an "ignorance inference" is something we will return to in §6.3. For now we can think of it as the inference that the speaker is ignorant about all alternatives that are not settled by her utterance — that she doesn't know whether both her children live in Europe nor whether they both live in the US nor whether some live in the US and some live in Europe. Ignorance inferences should be generated only when it is relevant where the children live; it seems reasonable that whenever sentences such as (4a) and (4b) are uttered the question where the children live would indeed be taken to be relevant. We think that this can be explained when thinking about possible QUDs, but we will not go into this for lack of space.

status of (3b) to favor the gradable generalization in (2)).

(5) **Categorical Denić generalization (CDG)**:
In sentences of the form $\forall x \in D(P_1 x \vee \ldots \vee P_n x)$, distributive inferences can only be derived if $n < |D|$.

Further evidence for the CDG comes from Denić herself, specifically from the explanation she provides for her observation of the contrasts in (6) and in (7).

(6) a. #Every one of these three girls is Sima, Rina or Dina.
b. Every one of these three girls is called Sima, Rina or Dina.

(7) a. (Tolstoy, Zola and Rowling are great writers.)
Each of those three writers wrote Anna Karenina, Germinal, or Harry Potter.
b. (Anne, Jane, and Bob are great students.)
Each of those three students read Anna Karenina, Germinal, or Harry Potter.

In all four sentences in (6) and (7), DIs are unavailable. This, Denić argues, leads to various ignorance inferences — that the speaker is ignorant about the DIs. However, for (6a) and (7a) ignorance of this sort is not compatible with basic facts anyone would take for granted in a conversation.[2] Specifically, it is impossible (at least by contextual knowledge) for two different girls to be Sima. So knowing that everyone of the students is either Sima, Rina or Dina (contextually) entails that one of the students is Sima, one of them is Rina and one of them is Dina. A speaker then cannot be taken to believe (6a) without believing its DIs. This contrasts with (6b), where is it entirely possible for two different girls to be *called* Sima. Likewise for (7a) and the contrast with (7b): if Author 1 wrote book 1, it follows that Author 2 did not, a consequence that does not hold for *read*, as two people can read the same book.[3]

We think this is a very plausible explanation, but Denić's generalization, as stated in (2), is not strong enough to derive it. Specifically, (2) leads us to expect that distributive inferences are going to be hard when $|D| - n = 0$ — harder than when $|D| - n = 1$ — but it does not lead us to expect the categorical judgment attested in (6a) and (7a). For the judgments to follow, we need to categorically block distributive inferences when $|D| = n$, not merely place them lower on a preference scale than comparable cases where $|D| - n$ is greater. This categorical blocking is precisely what the CDG aims to describe.

2 This characterization of course raises the question why these ignorance inferences should be obligatory (a question which Denić is aware of). See fn. 22.

3 An alternative perspective on the infelicity of (6a) and (7a) is provided by the Logical Integrity generalization proposed by Anvari (2018), according to which a sentence S is infelicitous if it has an alternative which is contextually but not logically entailed by S. Because (6a) and (7a) contextually but not logically entail their DIs (before implicature computation), Logical Integrity captures their infelicity. However, Logical Integrity does not expect the absence of distributive inferences when $n \geq |D|$ in general, and as a result cannot explain the categorical contrast in (4).

Admittedly, we could revise Denić's generalization more minimally. Specifically, we could keep the statement in (2) and add an auxiliary statement that the judgment becomes categorial when $|D| = n$ (that the dispreferred inference is categorically blocked). Our strategy will be different. We will begin with the weaker CDG and show that it follows from a particular way of looking at an independently needed ban on fatal competition. This account of the CDG will raise an obvious question about the gradient residue. But we will see in §8 that this question receives a rather simple answer based on fairly simple-minded ideas about the nature of disambiguation. In other words, we will end up with a picture in which the CDG follows from a modular theory of SIs, which is blind to probabilities, whereas the gradient effect follows from general assumptions about the nature of disambiguation, where probabilistic considerations are expected to figure prominently. But before we can get to any of this, we need to lay out a few background assumptions about the theory of SIs, as it is relevant for the derivation of DIs. In particular, we need to defend a proposal about the set of alternatives accessed in the derivation of DIs as they will be crucial for the emergence of fatal competition whenever $n \geq |D|$.

3 On the derivation of distributive inferences

It is well-known that the SIs of a sentence cannot be determined based on meaning alone. What is needed, in addition, is a specification of the set of alternatives associated with linguistic expressions.[4] So if we want to figure out the SIs of sentences such as (1), and, in particular, understand the conditions necessary for the derivation of DIs, we need to inquire into the nature of the alternatives associated with such sentences. With DIs in mind, it has often been assumed that the relevant sentences, schematized as $\forall x(Px \vee Qx)$, have the alternatives in figure 1 (Sauerland 2004a,b; Spector 2006; Fox 2007).

Assuming this set of alternatives, the sentence has a simple property which we can call universal excludability. By this we mean that it is consistent to assert the basic sentence $\forall x(Px \vee Qx)$, and to negate every one of the (distinct) alternatives. When universal excludability holds, all of the alternatives can be excluded without fear of contradiction, and theories of scalar implicatures therefore lead to their exclusion; the strengthened meaning of the sentence is thus as in (8).

(8) $\underbrace{\forall x(Px \vee Qx)}_{\text{Basic meaning}} \wedge \underbrace{\neg\forall x Px \wedge \neg\forall x Qx}_{\text{SIs}}$

This strengthened meaning entails the DIs. We can assume, if we'd like, a grammatical theory of

4 For that, one needs to look at the syntactic structure of the sentence as assumed already in Horn (1972), but very clearly articulated in Katzir (2007). There are many ways to see that this is the case, but one very obvious illustration comes from pairs of sentences that have identical basic meanings (before SIs are computed) and yet are associated with different SIs, e.g., *some students passed the test* and *some or all students passed the test*. See Chierchia et al. (2012) for an account.

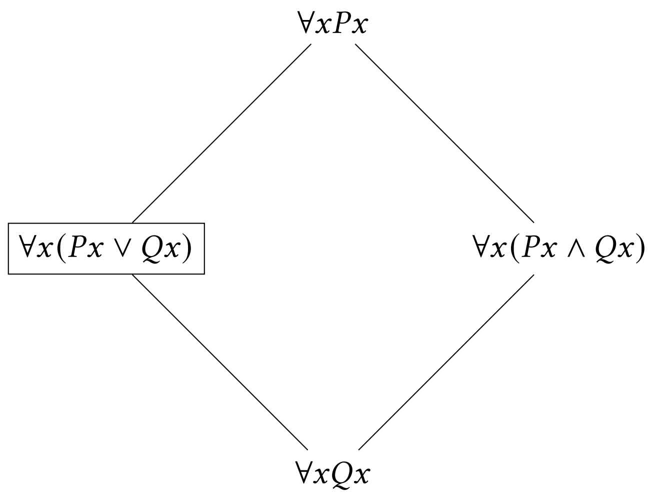

Figure 1 Restricted set of alternatives of $\forall x(Px \vee Qx)$

SIs, with the exhaustivity operator in (9) from Fox (2007), which negates Innocently Excludable (IE) alternatives. Letting $\mathscr{E}xh$ negate only IE alternatives (rather than, say, all stronger alternatives) is needed if we want to avoid deriving a contradiction in cases which do not have the property of universal excludability, such as unembedded disjunction (the relevance of these cases will shortly become clear). When applied to the basic sentence — the prejacent — and the set of alternatives in Figure 1, the result is what is described in (8). But for now we might also think of the exhaustivity operator as shorthand for pragmatic reasoning, as in van Rooij (2002). Adopting a grammatical theory will become crucial in a bit, as it will allow for recursive application of the operator, something that we will argue is necessary.[5]

(9) Definition of $\mathscr{E}xh$:

a. $[\![\mathscr{E}xh]\!](C)(p)(w) = 1$ *iff* $p(w) = 1 \wedge \forall q \in IE(C,p)[q(w) = 0]$

b. $IE(C,p) = \bigcap\{C' : C'$ is a maximal subset of C s.t. $p \wedge \bigwedge\{\neg q : q \in C'\} \not\Leftrightarrow_L \bot\}$

Whether a grammatical or a pragmatic account is adopted, the overall picture is rather simple. However, it is challenged by both theoretical and empirical observations. On the theoretical front, one needs to appeal to a general procedure that would determine the alternatives of a linguistic expression. If we adopt the structural approach to alternatives proposed in Katzir (2007) — or a proposal based on Horn Scales amended as in Sauerland (2004b) to yield the necessary alternatives for disjunctions — we would have not only the alternatives in Figure 1 but also the alternatives derived by substitution of the universal quantifier with its existential alternative, as in Figure 2. And with these alternatives, the sentence no longer has the property of universal excludability, and DIs are not derived. Specifically, exclusion of $\forall xPx$ entails (given the prejacent) $\exists xQx$, hence these

5 We use the notation $A \Rightarrow_L B$ when A logically entails B, and $A \Rightarrow_C B$ when A contextually entails B (i.e., when B is true in all worlds in the context set where A is true).

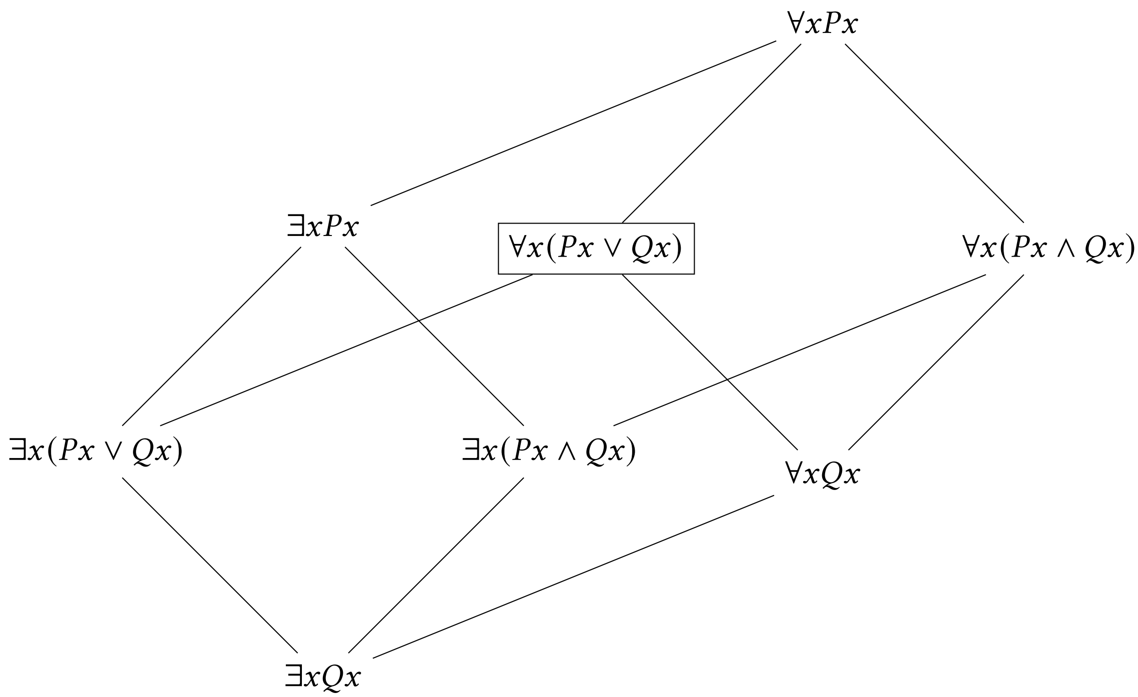

Figure 2 Full set of alternatives of $\forall x(Px \vee Qx)$

alternatives cannot be negated together while preserving consistency with the prejacent, and as a result they are not IE. For parallel reasons, $\forall xQx$ and $\exists xPx$ are not IE either. In fact, the only IE alternatives are $\forall x(Px \wedge Qx)$ and $\exists x(Px \wedge Qx)$ and DIs are thus not derived.

This theoretical observation is a serious challenge for a derivation of DIs on the basis of the alternatives in figure 1. To meet this challenge, Denić suggests that the set of alternatives in Figure 1 can be derived from the set in Figure 2 by the process of contextual-pruning, but that, we think, is highly problematic in light of the observation made in various places that pruning should never lead to a situation where a non-IE alternative becomes innocently excludable (a problem which Denić acknowledges). In other words, pruning should never "break symmetry" (see Fox and Katzir 2011; Katzir 2014; Crnič et al. 2015, as well as fn. 7 below).

There are two other responses to the challenge that have been suggested in the literature. In Fox (2007) it is suggested that alternatives, though formed by substitution, cannot be formed by substituting a strong alternative (e.g., a universal quantifier) with a weaker alternative (an existential quantifier) in an upward entailing environment. Another response is that alternatives of a sentence can be formed only by substitution of phrases that are contained in a focus marked constituent. If the subject quantifier in (1) is not contained in an F-marked constituent, the alternative set in Figure 1 would be formed and DIs would be derived.

While both responses are, in principle, reasonable, they are themselves challenged in other ways, and ultimately untenable, at least as far as we can see. The proposal in Fox (2007) has been challenged by observations that, in other cases, SIs depend on substitution with weaker alternatives (see Chemla 2009; Romoli 2012; Chemla and Romoli 2015; Bar-Lev and Fox 2020 for discussion). And the idea that distributive inferences rely on a particular focus placement is challenged by the observation of Chierchia (2004) about "intervention", namely his observation that alternatives generated in the scope of a DP cannot be accessed by an operator that c-commands the DP, unless the operator also accesses alternatives generated by the DP. So while F-marking has been shown to be crucial for the computation of alternatives, there are patterns of association with focus that are blocked in the syntax and those block an exhaustivity operator from having the alternatives in Figure 1.

But the clearest challenge to the idea that the alternatives in Figure 1 can be accessed by an exhaustivity operator comes directly from the derivation of DIs (Crnič et al. 2015), namely the observation that the predicted strengthened meaning is stronger than attested. Consider the contrast between (10a) and (10b).

(10) a. Every brother of mine is married to a man or a woman.
b. Every brother of mine has been married to a man or a woman.

(10a) is highly misleading if every brother of the speaker is married to a woman. This is presumably because in such a situation the DI that some brother is married to a man is false (assuming that it's impossible to be married to more than one person at the same time). (10b), by contrast, is entirely natural even if every brother has (at some point in time) been married to a woman. But the DIs still must hold: it is still required that one brother has (at some point in time) been married to a man and one (other) brother has (at some point in time) been married to a woman. This means that the DIs must follow without the negation of all of the alternatives in Figure 1, as we see in (11).

(11) Every brother of mine has been married to a man or a woman.
a. $\rightsquigarrow$ some brother of mine has been married to a man.
b. $\rightsquigarrow$ some brother of mine has been married to a woman.
c. $\not\rightsquigarrow$ it's not the case that every brother of mine has been married to a man.
d. $\not\rightsquigarrow$ it's not the case that every brother of mine has been married to a woman.

So, with the 4-membered alternative set in Figure 1, DIs are derived automatically, but the meaning is too strong and raises a bunch of theoretical questions as well. The 8-membered alternative set in Figure 2 is what we expect on theoretical grounds, but with it we appear to be unable to derive DIs at all.

This impasse is overcome, as pointed out in Bar-Lev and Fox (2016, 2020), the moment we allow the exhaustivity operator to apply recursively. With a second application of $\mathcal{E}xh$ we do derive DIs: On the first application of $\mathcal{E}xh$ the conjunctive alternatives $\forall x(Px \wedge Qx)$ and $\exists x(Px \wedge Qx)$ are excludable; on the second application of $\mathcal{E}xh$ two alternatives are excludable: $\mathcal{E}xh(\forall x Px)$ and $\mathcal{E}xh(\forall x Qx)$ (other alternatives are excludable as well, but their exclusion is vacuous). The exclusion of these exhaustified alternatives yields distributive inferences:

(12) a. $\mathscr{E}xh(\forall x Px) = \forall x Px \wedge \neg\exists x Qx$

b. $\mathscr{E}xh(\forall x Qx) = \forall x Qx \wedge \neg\exists x Px$

(13) $\mathscr{E}xh\mathscr{E}xh(\forall x(Px \vee Qx))$

$= \forall x(Px \vee Qx) \wedge \neg\mathscr{E}xh(\forall x Px) \wedge \neg\mathscr{E}xh(\forall x Qx) \wedge \neg\exists x(Px \wedge Qx)$

$= \forall x(Px \vee Qx) \wedge \neg(\forall x Px \wedge \neg\exists x Qx) \wedge \neg(\forall x Qx \wedge \neg\exists x Px) \wedge \neg\exists x(Px \wedge Qx)$

$= \forall x(Px \vee Qx) \wedge \exists x Px \wedge \exists x Qx \wedge \neg\exists x(Px \wedge Qx)$

Moreover, the results are compatible with Crnič et al. (2015). Specifically, we are allowed to prune the existential conjunctive alternative $\exists x(Px \wedge Qx)$ — without thereby breaking any symmetry — in which case the meaning is compatible with the truth of $\forall x Px$ or $\forall x Qx$.[6,7]

So we have seen that recursive application of $\mathscr{E}xh$ allows us to derive DIs from the alternative set in Figure 2 with better empirical results than what can be achieved with the theoretically problematic set in Figure 1. Our goal is now to show that the alternative set in Figure 2 also provides an account of the CDG based on independent observations pertaining to fatal competition. But before that we still have to go over the derivation of Free Choice in Fox (2007). This derivation is entirely parallel to the derivation of DIs above, thus providing further support for it. But, more importantly, it is crucial for assessing a property of one of the eight alternatives in Figure 2, precisely the property that under certain circumstances can lead to fatal competition.

4 Free Choice

The conjunctive inferences attested when disjunction is embedded under an existential modal, as in (14), are surprising given the basic semantics predicted for the construction (von Wright 1968; Kamp 1974; Kratzer and Shimoyama 2002, among many others).

(14) You may read Anna Karenina or The Brothers Karamazov. $\Diamond(p \vee q)$

a. $\rightsquigarrow$ You may read Anna Karenina $\Diamond p$

b. $\rightsquigarrow$ You may read The Brothers Karamazov $\Diamond q$

6 Crnič et al. (2015) propose a way to derive DIs without the negation of $\forall x Px$ and of $\forall x Qx$. Their derivation is however dependent on the stipulation that the existential alternatives are not formal ones (or that they are pruned), which, as we have argued above, is difficult to reconcile with facts about intervention effects (or symmetry breaking). We achieve the same result without the stipulation. Furthermore, as we will discuss in the following sections, our explanation for the CDG will crucially rely on the assumption that existential alternatives are necessary for the derivation of DIs; this account will then be incompatible with Crnič et al.'s proposal.

7 The possibility of deriving DIs by recursive $\mathscr{E}xh$ given the set of alternatives in Figure 2 is problematic for Denić's view where the set of alternatives in Figure 2 can only lead to ignorance inferences about the DIs (and the derivation of DIs is contingent on having the restricted set of alternatives in Figure 1). In order to maintain the picture she argues for, then, recursive exhaustification would have to be blocked.

The inferences — referred to as Free Choice inferences, FCIs — do not follow from the basic meaning of the sentence, and like DIs have been argued to follow as SIs (Kratzer and Shimoyama 2002; Alonso-Ovalle 2005). The derivation of FCIs in Fox (2007) is based on recursive exhaustification and is completely parallel to the derivation of DIs presented in §3.[8] To see the similarity it will be helpful to pretend that the existential modals in (14) are existential nominal quantifiers and thereby look at the set of alternatives we already looked at (in Figure 2). This pretense is completely innocuous as every alternative will have a parallel in the modal domain, with all logical relations maintained. We will see that although the derivation starts with a different prejacent it ends up having completely parallel steps and virtually identical results.

Before we run through the derivation it is useful to see the similarity in a first pass description of the results: both FCIs and DIs are the inferences that the simple existential alternatives $\exists xPx$ and $\exists xQx$ are true. Now to the derivation of FCIs and a more specific description of the similarity. On the first application of $\mathcal{E}xh$ the conjunctive alternative $\exists x(Px \wedge Qx)$ is excludable, as well as all universal alternatives (which amounts to the exclusion of $\forall x(Px \vee Qx)$). On the second application of $\mathcal{E}xh$ two alternatives are excludable: $\mathcal{E}xh(\exists xPx)$ and $\mathcal{E}xh(\exists xQx)$.

(15) a. $\mathcal{E}xh(\exists xPx) = \exists xPx \wedge \neg\exists xQx$
b. $\mathcal{E}xh(\exists xQx) = \exists xQx \wedge \neg\exists xPx$

(16) $\mathcal{E}xh\mathcal{E}xh(\exists x(Px \vee Qx))$
$= \exists x(Px \vee Qx) \wedge \neg\mathcal{E}xh(\exists xPx) \wedge \neg\mathcal{E}xh(\exists xQx) \wedge \neg\exists x(Px \wedge Qx) \wedge \neg\forall x(Px \vee Qx)$
$= \exists x(Px \vee Qx) \wedge \neg(\exists xPx \wedge \neg\exists xQx) \wedge \neg(\exists xQx \wedge \neg\exists xPx) \wedge \neg\exists x(Px \wedge Qx) \wedge \neg\forall x(Px \vee Qx)$
$= \exists x(Px \vee Qx) \wedge \exists xPx \wedge \exists xQx \wedge \neg\exists x(Px \wedge Qx) \wedge \neg\forall x(Px \vee Qx)$

The results we get for $\forall x(Px \vee Qx)$ and for $\exists x(Px \vee Qx)$ are almost identical: with both a universal prejacent and an existential prejacent we derive the inferences $\exists xPx \wedge \exists xQx \wedge \neg\exists x(Px \wedge Qx)$. The only difference between the exhaustified meaning in the two cases is whether $\forall x(Px \vee Qx)$ is entailed or its negation is. That is, the underlined parts in (17) are equivalent, and the only difference is between the boldfaced parts.

(17) a. $\mathcal{E}xh\mathcal{E}xh(\forall x(Px \vee Qx)) = \boldsymbol{\forall x(Px \vee Qx)} \wedge \underline{\exists xPx \wedge \exists xQx \wedge \neg\exists x(Px \wedge Qx)}$
b. $\mathcal{E}xh\mathcal{E}xh(\exists x(Px \vee Qx)) = \underline{\exists x(Px \vee Qx) \wedge \exists xPx \wedge \exists xQx \wedge \neg\exists x(Px \wedge Qx)} \wedge \boldsymbol{\neg\forall x(Px \vee Qx)}$

This fact will play a key role in our explanation of the CDG. It will also be crucial to assume that FCIs arise not only with existential modals but also with existential quantifiers. This latter assumption,

8 Bar-Lev and Fox (2020) have argued that free choice inferences are derived through an Innocent Inclusion procedure (part of an alternative definition of $\mathcal{E}xh$). This makes the application of a second $\mathcal{E}xh$ operator redundant for the derivation of free choice, but not for the derivation of distributive inferences. Since our main focus here is distributive inferences we do not rely on Innocent Inclusion, though our proposal is compatible with it. If we adopt Innocent Inclusion, recursive exhaustification can still apply and its consequences will be exactly identical to what one gets without Innocent Inclusion where distributive inferences are involved.

although defended in the literature, might raise questions, as the relevant inferences are more scarce in the nominal domain. Nevertheless, evidence for their presence has been presented in the literature along with an explanation for the limited distribution (Fox 2007; Klinedinst 2007; Eckardt 2007):

(18) a. Some took the highway or the scenic route.
b. (If you want a drink,) there is beer in the fridge or the ice bucket.
c. There is cake or ice cream in the refrigerator (in case you are hungry for dessert).

Given this, we will assume that the inferences are formally available (with non-singular existential quantifiers), an assumption that will, again, be crucial for our explanation of the CDG.

5 Fatal competition

5.1 Key observation: equivalence when $n \geq |D|$

Let us start by looking at what happens when we consider example (3c) where $n = |D|$ (repeated in (19)), and its existential alternative in (20). Suppose we compute all potential SIs for (19) and (20) — DIs and FCIs respectively — ignoring, for now, the SI of (20) that pertains to (19), namely the exclusion of (19) by $\mathcal{E}xh$((20)).

(19) Both of Mary's friends are French or Spanish
a. **Strong exclusive inference**: None of Mary's friends are both French and Spanish
b. **Distributive inferences**:
(i) At least one of Mary's friends is French
(ii) At least one of Mary's friends is Spanish

(20) Some of Mary's friends are French or Spanish
a. **Strong exclusive inference**: None of Mary's friends are both French and Spanish
b. **Free Choice inferences**:
(i) At least one of Mary's friends is French
(ii) At least one of Mary's friends is Spanish

When taken together with their SIs (again excluding inferences about each other), (19) and (20) turn out to be equivalent. The implication from (19) to (20) is trivial ($\forall \Rightarrow \exists$). To see that the implication goes in the other direction, consider what is needed for (20) to be true. There have to be two friends, call them Friend 1 and Friend 2, such that Friend 1 is Spanish (and not French) and Friend 2 is French (and not Spanish):

	Friend 1	Friend 2
French	✓	✗
Spanish	✗	✓

But since there are only two friends in the domain, this means that (19) is true as well. So (19) and (20) are equivalent when $|D| = n = 2$. And it is easy to see that the equivalence breaks down the moment we enlarge the quantification domain, D, and keep the number of disjuncts constant: if neither disjunct is true of the the third individual in the domain, the existential sentence would be true and the universal sentence would be false:

	Friend 1	Friend 2	Friend 3
French	✓	✗	✗
Spanish	✗	✓	✗

And of course, it is easy to see that if we add a third disjunct (and a third DI, FCI) the equivalence would re-emerge. More generally, (21a) and (21b) are equivalent iff $n \geq |D|$.[9]

(21) a. $\forall x \in D(P_1x \vee \ldots \vee P_nx) \wedge \bigwedge_{1\leq i\leq n} \exists x \in D(P_ix) \wedge \bigwedge_{\substack{1\leq i,j\leq n\\ i\neq j}} \neg\exists x \in D(P_ix \wedge P_jx)$

b. $\exists x \in D(P_1x \vee \ldots \vee P_nx) \wedge \bigwedge_{1\leq i\leq n} \exists x \in D(P_ix) \wedge \bigwedge_{\substack{1\leq i,j\leq n\\ i\neq j}} \neg\exists x \in D(P_ix \wedge P_jx)$

This equivalence, we propose, is key to understanding why DIs cannot be derived when $n \geq |D|$: They would result in fatal competition, a notion that we will exemplify in the next subsection and characterize more precisely in §6, relying heavily in both cases on the work of Magri (2009a,b), Spector (2014), and Anvari (2018). But to understand what's on the horizon, we suggest here that a situation of fatal competition arises whenever there are existential and universal alternatives which have the properties in (22).

(22) **Characterization of (certain) cases of fatal competition**:

a. One member of the set of alternatives (e.g., $\exists x\phi$) is contextually equivalent to a logically stronger alternative (e.g., $\forall x\phi$) when the exclusion of the stronger alternative is ignored.[10]

9 The entailment from (21a) to (21b) is again trivial. To show that the entailment goes in the other direction when $n \geq |D|$, it is enough to show that in such a case the conjunction of (ia) and (ib) entails (ic).

(i) a. $\bigwedge_{1\leq i\leq n} \exists x \in D(P_ix)$

b. $\bigwedge_{\substack{1\leq i,j\leq n\\ i\neq j}} \neg\exists x \in D(P_ix \wedge P_jx)$

c. $\forall x \in D(P_1x \vee \ldots \vee P_nx)$

Note that (ia) and (ib) together entail that for every predicate P_i, there is an individual in D which satisfies P_i and satisfies no other predicate. From this it follows that there are at least n individuals x in D for which $P_1x \vee \ldots \vee P_nx$ is true. Now if (and only if) $n \geq |D|$, (ic) follows. The same holds if (ib) is replaced with the weaker Distinctness assumption we will introduce in Appendix A.

b. Subsequently, the weaker alternative is contradictory when the exclusion of the stronger alternative by the weaker one is taken into account.

As we have just shown, (22a) holds of sentences of the form $\forall x \in D(P_1 x \vee \ldots \vee P_n x)$ when $n \geq |D|$ and DIs are obtained, and consequently, (22b) holds as well. This observation, we propose, provides an account of the CDG. We note also that the account crucially depends on the assumptions we defended above regarding the derivation of DIs, namely that existential alternatives are crucial and that the set in Figure 1 cannot be the entire set of alternatives nor can it be derived from the formal set via contextual pruning. On views where DIs are derived with sets of alternatives such as that in Figure 1 (i.e., without any existential sentences as alternatives, e.g., Fox 2007; Crnič et al. 2015; Denić 2021), the equivalent existential sentence would not be in the set of alternatives and hence could not lead to fatal competition. In the next subsection we discuss other well-known cases of fatal competition and show that they share the property in (22).

5.2 Other known cases of fatal competition

The idea that alternatives resulting from substitution of logically distinct elements are deviant when contextually equivalent is not new. A pattern that has been discussed extensively pertains specifically to existential and universal quantification. Consider the pair in (23) (see Magri 2009a,b, 2011; Spector 2014; Anvari 2018, a.o.):

(23) a. #All Italians come from a warm country.
b. #Some Italians come from a warm country.

The two sentences are logically distinct, as it is logically possible (as far as the meaning of the logical words is concerned) for some Italians to come from a warm country and for others not to. But the two sentences are equivalent once the meaning of the non-logical words (or our beliefs pertaining to them) are factored in, once it is accepted (as common ground) that Italians all come from the same country.[11] Of course the equivalence only holds if we don't factor in the SI of (23b) which involves the negation of (23a). In other words, the two alternatives in (23) exemplify the properties of fatal competition which we just discussed.[12]

10 To understand this characterization of fatal competition we, of course, need to understand what it means for two sentences to be equivalent once we ignore the exclusion of one of them by the other. We think the notion is clear enough but here is one way things can be stated:

(i) An alternative A of S is contextually equivalent to S when inferences from A about S are ignored *iff* $\mathscr{E}xh_{Alt(S)}(S) \Leftrightarrow_C \mathscr{E}xh_{Alt(S)\setminus\{S\}}(A)$.

11 Magri (2011) discusses cases similar to (23) where the (a) variant is acceptable. We hope that there is something principled that could be said about these cases that would be compatible with our proposal; see Singh (2009); Spector (2014) for relevant ideas, observations and further challenges.

12 The accounts entertained in the literature for these cases of fatal competition presupposes that (23a) and (23b) are necessarily alternatives of each other. We note that this is not the case under all theories of alternatives. For example,

This similarity between well-known cases of fatal competition, such as (23), and cases where DIs are ruled out by the CDG (that is, the fact that both cases share the fatal competition properties in (22)), suggests that they should be captured uniformly. But unfortunately none of the current accounts can be appealed to: Magri (2009a,b) and Anvari (2018) utilize the property in (22b) to explain the oddness of (23b) (taking the oddness of (23a) to come from a different source), but their proposals do not account for the CDG. Spector (2014) utilizes the property in (22a) to explain both the oddness of (23a) and (23b), but his proposal also does not account for the CDG. We thus need to develop an alternative, which is what we attempt to do in the next section, building on Fox (2018) and more directly on unpublished work by Amir Anvari.

6 Proposal

6.1 The general proposal

It is commonly assumed that an indicative sentence must be interpreted as an answer to a question. Very often this question is not present in the preceding discourse and needs to be inferred/accommodated. But in any event, the general idea imposes an obvious condition on acceptability: For an indicative sentence S to be acceptable, there must be a good question, Q, that can be accommodated, such that S could be interpreted as a good answer to Q. This general principle, articulated for example in Katzir and Singh (2015), is what we might call a bare output condition, extending the term from Chomsky (1995).

The condition, however, can be strengthened if we assume, following, e.g., Rooth (1992), that the retrievable question must bear a particular relationship to the formal alternatives induced by the answer, i.e., its focus alternatives. But we will consider an additional condition entertained by Amir Anvari as a theory of fatal competition and related to a proposal in Fox (2018), namely that the relevant set of alternatives must partition the context set.

(24) **Sentence-Question Correspondence**:

a. A sentence S is felicitous only if there is a question Q which is a good question for S.

b. Q is a good question for S only if

(i) $Q \supseteq Alt(S)$.

(ii) Q partitions the context set C.[13]

(25) Q partitions C if all of the following conditions hold:

Fox and Katzir (2011) argue that alternatives can only be derived by substitutions within a focused phrase. If this is the case, we would expect the sentences to improve when the subject is not dominated by a focus-marked constituent. We are not sure that this is the case (though see Singh 2009). In any event, the issue does not arise for our account of the CDG. For that account it is sufficient to look at the set of alternatives that are needed for deriving DIs and this, as we argue in §3, requires the entire set of alternatives in Figure 2.

a. **Non-Emptiness**: For all $S \in Q$: $[\![S]\!] \cap C$ is not the empty set.
b. **Disjointness**: For all $S1, S2 \in Q$: $[\![S1]\!] \cap C$ and $[\![S2]\!] \cap C$ are disjoint.
c. **Completeness**: $\bigcup\{[\![S]\!] \cap C : S \in Q\} = C$

For expository purposes, we assume that Q must contain all of S's formal alternatives (as in (24b-i)). In Appendix A we consider the implications of a more permissive assumption, namely that Q must contain a possibly smaller set derived by a process of pruning from the set of formal alternatives.[14] In order to see the consequences of (24), it will be helpful to focus on the following requirement, which turns out to be equivalent to (24):[15]

(26) A sentence S is felicitous only if there is a subset C' of the context-set C s.t. $Alt(S)$ partitions C'.

In many cases, satisfying this condition will require applying $\mathcal{E}xh$. Consider for instance the utterance of (27a) with its alternative in (27b).

(27) a. Some students smiled.
b. All students smiled.

The two alternatives cannot partition any subset C' of the context-set C, simply because (27b) entails (27a). Consider $[\![(27b)]\!] \cap C'$: If it's not empty, then there is a $w \in [\![(27b)]\!] \cap C'$, which, given entailment, is also in $[\![(27a)]\!] \cap C'$, and hence disjointness is not satisfied. If $[\![(27b)]\!] \cap C'$ is empty, then non-emptiness isn't satisfied.

But if we add $\mathcal{E}xh$ to the parse of (27a), we get the two alternatives in (28) where entailment no longer holds. Moreover, the two sentences now contradict each other and hence can partition some subset of C.[16,17]

(28) a. $\mathcal{E}xh$ Some students smiled (= Some but not all students smiled).
b. $\mathcal{E}xh$ All students smiled (= All students smiled).

13 The reason we do not identify Q with $Alt(S)$ (and merely require Q to be a superset of $Alt(S)$) is due to the fact that, in many cases, a set of alternatives doesn't cover all worlds in the context set. For instance, *some kids smiled* and *all kids smiled* have no alternative which is true in worlds where no kid smiled (regardless of whether $\mathcal{E}xh$ applies).

14 This is of course related to the fact that Rooth's requirement demands that the focus value of S be a super set of the congruent question, which might be restated as the demand that the focus value of the sentence, after pruning, be identical to the question.

15 Fulfilling the requirement in (26) is both necessary and sufficient for fulfilling the one in (24). It is necessary because if Q partitions C and $Alt(S)$ is a subset of Q, then $Alt(S)$ must partition some subset of C. It is sufficient because if $Alt(S)$ is a partition of some subset of C, then we can construct Q by adding a sentence to $Alt(S)$ namely the negation of all sentences in $Alt(S)$; constructed this way, Q is guaranteed to partition C.

16 In some cases $\mathcal{E}xh$ alone is not enough to derive a partition; see our discussion in §6.3.

17 This appears to make the wrong prediction that the scalar implicature for *some* will be obligatory, as we've abstracted away from pruning. This apparent prediction can be avoided independently of assumptions about pruning if we move to partition by $\mathcal{E}xh$ along the lines of Fox (2018). See Appendix A.

Let us now turn to the consequences of this proposal for simple cases of fatal competition. It predicts unacceptability for cases where *some* and *all* sentences are contextually equivalent, such as (23), because, whether or not $\mathscr{E}xh$ applies, $Alt(S)$ does not partition any subset of the context set: If $\mathscr{E}xh$ doesn't apply, then we have two equivalent alternatives in $Alt(S)$, and as a result disjointness cannot be satisfied. If $\mathscr{E}xh$ does apply, we have an alternative in $Alt(S)$ which is a contextual contradiction (since $\mathscr{E}xh(some) =$ *some but not all*); as a result non-emptiness is violated (the idea that a question with contextually contradictory cells is not allowed has precedent in the maxim of Maintain Uniformity proposed in Singh 2009).[18] Now we can see how this proposal also predicts the more complex case of fatal competition we are interested in, that is how it derives the CDG.

6.2 Deriving the CDG

Recall from §3 that in order to derive DIs we need two applications of $\mathscr{E}xh$ to a sentence of the form $\forall x(Px \vee Qx)$, as in (29a). The sentence will then have the alternative in in (29b). As we established, (29a) and (29b) are equivalent if $2 \geq |D|$ once we ignore the inference from the existential alternative that the universal one is false (that is if we remove the last conjunct in (29b)).

(29) a. $\mathscr{E}xh\mathscr{E}xh(\forall x(Px \vee Qx)) = \boldsymbol{\forall x(Px \vee Qx)} \wedge \underline{\exists xPx \wedge \exists xQx \wedge \neg\exists x(Px \wedge Qx)}$

b. $\mathscr{E}xh\mathscr{E}xh(\exists x(Px \vee Qx)) = \underline{\exists x(Px \vee Qx) \wedge \exists xPx \wedge \exists xQx \wedge \neg\exists x(Px \wedge Qx)} \wedge \boldsymbol{\neg\forall x(Px \vee Qx)}$

As a result, when this inference is not ignored (as in (29b)), we get a set of alternatives one of which is a contextual contradiction when $2 \geq |D|$, a set of alternatives which cannot partition any set of worlds. More generally, it follows that:

(30) $Alt(\mathscr{E}xh\mathscr{E}xh(\forall x \in D(P_1x \vee \ldots \vee P_nx)))$ is not a partition of a subset of the context set if $n \geq |D|$.

The derivation of distributive inferences is blocked if $n \geq |D|$ given (26) and (30), due to the perspective on DIs defended in §3.

We thus have a unified account of cases where DIs are ruled out by the CDG and cases of 'oddness' discussed by Magri/Spector, based on the observation that in both varieties exhaustification renders one of the alternatives a contextual contradiction. There is, however, an empirical distinction between the Magri/Spector cases and the CDG that we need to explain: The Magri/Spector cases

18 Having contextual contradictions as alternatives doesn't always lead to oddness (Roni Katzir, p.c.):

(i) Mary has five broken fingers on her right hand.
Contextually contradictory alternative: Mary has six broken fingers on her right hand.

The issue does not arise for Spector (2014), and requires a weakening of our constraint. In Appendix A we suggest that contextual contradictions like the one in (i), which unlike other contextual contradictions we discuss in this paper are not the result of exhaustification, can be ignored by a process of alternative pruning.

are simply unacceptable while the cases where DIs are ruled out by the CDG are not. When CDG rules out DIs, we get ignorance inferences rather than unacceptability, as argued by Denić (2021). As it stands, however, our proposal makes the wrong prediction that cases where DIs are not derived should be unacceptable (that is, regardless of the relation between n and $|D|$).

In the next section we show that we can avoid this wrong prediction and account for the difference between the Magri/Spector cases and cases where DIs are ruled out by the CDG once we adopt the grammatical theory of ignorance inferences proposed in Meyer (2013), something we need to do on independent grounds.[19] Our explanation will be based on a crucial difference between the two cases, which pertains to the nature of the alternative set before (iterative) exhaustification applies: Whereas in the Magri/Spector cases the existential and universal alternatives are equivalent from the get go, in the cases where DIs are ruled out by the CDG they are not: only when DIs are derived do the existential and universal alternatives become equivalent.

6.3 Ignorance inferences and Sentence-Question Correspondence

Before we discuss how the derivation of ignorance inferences in Meyer's theory avoids the wrong prediction that sentences should be unacceptable whenever the CDG blocks DIs, let us consider first what we predict for simple disjunction as in (31). As we already know by now, without $\mathcal{E}xh$ application the set of alternatives cannot partition any subset of the context set because (31) is entailed by its conjunctive alternative. As above, we can try to apply $\mathcal{E}xh$ in order to derive a partition, but in this case it will not suffice: Since (32a) is the disjunction of (32b) and (32c), a set of alternatives containing all of these alternatives cannot be a partition.

(31) Mary or Sue smiled

(32) a. $\mathcal{E}xh$ Mary or Sue smiled (= Mary or Sue smiled but not both)
b. $\mathcal{E}xh$ Mary smiled (= Mary but not Sue smiled)
c. $\mathcal{E}xh$ Sue smiled (= Sue but not Mary smiled)
d. $\mathcal{E}xh$ Mary and Sue smiled (= Mary and Sue smiled)

There is however a way to derive a partition in this case once we adopt Meyer's (2013) theory of ignorance inferences, according to which they are derived by applying $\mathcal{E}xh$ above a universal K(nowledge) operator, which is a universal modal paraphrasable as 'the speaker is certain that...' (see Fox 2016; Buccola and Haida 2019 for conceptual grounding under Gricean assumptions). As (33) demonstrates, this structure gives rise to ignorance inferences and to a partition.[20]

19 This wrong prediction can also be avoided by pruning or by a move to partition by $\mathcal{E}xh$, as we will discuss in Appendix A.

20 There might also be cases which require pruning of all of the disjunctive alternatives, a possibility compatible with our proposal in Appendix A.

(33) a. $\mathscr{E}xh$ K $\mathscr{E}xh$ Mary or Sue smiled (= The speaker is certain that Mary or Sue smiled but not both, and she is not certain that Mary smiled and she is not certain that Sue smiled)
b. $\mathscr{E}xh$ K $\mathscr{E}xh$ Mary smiled (= The speaker is certain that Mary but not Sue smiled)
c. $\mathscr{E}xh$ K $\mathscr{E}xh$ Sue smiled (= The speaker is certain that Sue but not Mary smiled)
d. $\mathscr{E}xh$ K $\mathscr{E}xh$ Mary and Sue smiled (= The speaker is certain that Mary and Sue smiled)

This derivation of ignorance inferences solves the problem identified in the previous sub-section: it allows us to have a partition without deriving DIs for sentences of the form $\forall x(Px \vee Qx)$, thereby avoiding the wrong prediction we made in the absence of a K operator. In this case too, we derive ignorance inferences by applying $\mathscr{E}xh$ above K, and as a result we get a set of alternatives which can partition a subset of the context set (to avoid clutter we only compare here the universal and existential alternatives):[21]

(34) a. $\mathscr{E}xh(K(\forall x(Px \vee Qx))) = K(\forall x(Px \vee Qx)) \wedge \neg K(\exists x Px) \wedge \neg K(\exists x Qx)$
b. $\mathscr{E}xh(K(\exists x(Px \vee Qx))) = K(\exists x(Px \vee Qx)) \wedge \neg K(\exists x Px) \wedge \neg K(\exists x Qx) \wedge \neg K(\forall x(Px \vee Qx))$

Importantly, however, exhaustifying above K cannot rescue cases such as (23), where two alternatives are equivalent before exhaustification, since if $p \Leftrightarrow q$, then $Kp \wedge \neg Kq$ is a contradiction, just like $p \wedge \neg q$. The constraint on Sentence-Question Correspondence in (24), then, provides us with a unified account of the CDG and the Magri/Spector cases, and at the same time can also explain (with the aid of Meyer's theory of ignorance inferences) why the Magri/Spector cases are simply unacceptable, whereas for sentences of the form $\forall x \in D(P_1 x \vee \ldots \vee P_n x)$ distributive inferences are barred when $n \geq |D|$, but they are otherwise acceptable as long as ignorance inferences are derived. It also explains Denić's observation from §2 that cases where DIs are ruled out by the CDG are odd when ignorance inferences are incompatible with the common ground, under the assumption that it's impossible to derive neither distributive inferences nor ignorance inferences.[22]

7 Further predictions

Our explanation for the CDG was based on the observation that distributive inferences lead to fatal competition when $n \geq |D|$. In this section we discuss other cases where DIs are predicted to lead to fatal competition. We will first focus in §7.1 on a stronger generalization than the CDG, one where a larger domain of individuals than dictated by the CDG can bring about fatal competition: The governing factor, we will see, is not the number of individuals in the domain, but rather the number

21 Applying a single $\mathscr{E}xh$ operator below K as well will also derive a partition.

22 This assumption in fact follows from Sentence-Question Correspondence as stated in (24); things become more complicated if Sentence-Question Correspondence applies to the output of pruning rather than to the set of formal alternatives, as we entertain in (46) (though as we discuss in Appendix A, assuming that pruning affects Question-Sentence correspondence is not at all necessary).

of equivalence classes determined with reference to properties of the relevant disjuncts. We will then discuss in §7.2 cases where fatal competition arises regardless of domain size.

7.1 When domain-size is measured by the cardinality of equivalence classes

The following generalization follows from our proposal.[23,24]

(35) a. A sentence of the form $\forall x \in D(P_1 x \vee \ldots \vee P_n x)$ can have distributive inferences only if $m > n$ when m is the number of equivalence classes given by the equivalence relation R on D.

b. $R = \{\langle a, b\rangle \in D : \text{for all } P_i, P_i a \text{ is contextually equivalent to } P_i b\}$

Note first that (35) is stronger than the CDG: Because the number of equivalence classes m cannot be larger than the domain size $|D|$, whenever $n \geq |D|$ it must be the case that $n \geq m$. As a result, (35) covers all the cases where DIs are ruled out by the CDG. (35), however, rules out DIs in other cases as well. Let us demonstrate this by looking at the examples in (36), which are all modified versions of the Magri/Spector case in (23a). The CDG has nothing to say about any of the examples in (36) because in all of them $|D|$ is much greater than the number of disjuncts; Our proposal, however, since it has (35) as consequence, makes intricate predictions here.

(36) a. #Every Italian woman comes from a warm country.

b. Every French or Italian woman comes from a warm country.

c. Every French or Italian woman comes from a warm or beautiful country.

---

23 As far as we can tell, Denić (2021) makes a parallel gradable prediction.

24 In order to demonstrate why (35) is predicted, we need to show that whenever $m \leq n$ the universal alternative (ii) follows from the distributive inferences in (ia) together with the strong exclusive inferences in (ib). As we established, whenever this entailment goes through we predict distributive inferences to be blocked.

(i) a. $\bigwedge_{1 \leq i \leq n} \exists x \in D(P_i x)$

b. $\bigwedge_{\substack{1 \leq i,j \leq n \\ i \neq j}} \neg \exists x \in D(P_i x \wedge P_j x)$

(ii) $\forall x \in D(P_1 x \vee \ldots \vee P_n x)$

Consider two arbitrary disjuncts P_i and P_j (where $1 \leq i, j \leq n$). Given the distributive inferences in (ia) together with the exclusive inferences in (ib), there must be at least one individual that satisfies P_i but not P_j (and vice versa). In other words, there can be no two predicates P_i, P_j such that the set of R-based equivalence classes containing individuals who satisfy P_i is the same as the set of R-based equivalence classes of individuals who satisfy P_j. As a result, the number m' of R-based equivalence classes with individuals who satisfy at least one predicate is greater than or equal to the number of predicates n (i.e., $n \leq m'$). And of course the number m' of R-based equivalence classes with individuals who satisfy at least one predicate is smaller than or equal to the number of all R-based equivalence classes (i.e., $m' \leq m$). Now if the number of predicates n is greater than or equal to the number of all R-based equivalence classes m (i.e., $m \leq n$), it follows that $m = m'$ (because $n \leq m'$, $m' \leq m$, and $m \leq n$). That is, every individual in every R-based equivalence class must satisfy at least one predicate. And if that's the case, the universal alternative in (ii) follows. The same holds if (ib) is replaced with the weaker Distinctness assumption we will introduce in Appendix A.

(no distributive inferences)

(36a), just like (23a), is a case where a universal sentence is equivalent to its existential alternative. Since in (36a) there is only one disjunct and one equivalence class (for any pair of Italian women a and b, a comes from a warm country iff b does), our proposal predicts distributive inferences to be ruled out because $n = m$. A distributive inference when there is just one disjunct is simply the inference that the existential alternative is true, so the ban on distributive inferences simply means that (36a) cannot have as inference that *Some Italian woman comes from a warm country*. But since this inference is a matter of logical entailment, the sentence is predicted to be unacceptable. In (36b) n no longer equals m: We still have one disjunct but the number of equivalence classes is now 2 (for any pair of Italian women a and b, a comes from a warm country iff b does, and for any pair of French women a and b, a comes from a warm country iff b does). Consequently, distributive inferences, which are again logical entailments, are predicted to be possible. In (36c) n equals m again: Here we have two disjuncts and two equivalence classes, so distributive inferences are predicted to be impossible. In this case, however, they are not logical entailments so the sentence is not predicted to be unacceptable if ignorance inferences are derived. That is, our prediction is that (36c) can only be understood with ignorance inferences, namely that the speaker is not certain that France or Italy are warm and is not certain that either is beautiful. While we think that this prediction is borne out, it is also pretty clear that it is not easily available to introspection.

The following is a manipulation that we hope might yield sharper judgments. We construct a situation where the speaker is assumed to be completely knowledgeable about the relevant alternatives. Recall from our discussion of (4) that in such a situation the result is odd if distributive inferences are blocked (presumably because ignorance inferences are then derived, which contradict the assumption of complete knowledgeability). In (37) there are 4 children in the domain, but since the context entails that two of them live together we have 3 equivalence classes. Having 3 equivalence classes and 2 disjuncts ($m > n$), distributive inferences are predicted to be possible.

(37) A: I am sad that I can't see my **four** children on a regular basis. They live so far away.
B: Do they live alone?
A: No, in fact **Sally and Jane are roommates**. (The two others have partners.)
B: That's nice. Where do the four children live?
A: They each live in Europe or the United States.

In (38), by contrast, there are 3 children in the domain so since the context entails that two of them live together we have 2 equivalence classes. Similarly in (39), there are 4 children in the domain so since the context entails that three of them live together we have 2 equivalence classes. Having 2 equivalence classes and 2 disjuncts ($m = n$) in both cases, distributive inferences are predicted to be impossible. And given that ignorance inferences are derived in the absence of distributive ones, the result is predicted to be odd because we expect a person to know where their children live. These seem to us to be correct predictions.

(38) A: I am sad that I can't see my **three** children on a regular basis. They live so far away.
B: Do they live alone?
A: No, in fact **Sally and Jane are roommates**. (Mark has a partner)
B: That's nice. Where do the three children live?
A: #They each live in Europe or the United States.

(39) A: I am sad that I can't see my **four** children on a regular basis. They live so far away.
B: Do they live alone?
A: No, in fact **Sally, Jane and Mark are roommates**. (Bill has a partner)
B: That's nice. Where do the four children live?
A: #They each live in Europe or the United States.

7.2 When domain size is not relevant

Another good prediction we make concerns the unacceptability of (40) (an example due to Brian Buccola, p.c.). The unacceptability of (40) may seem unsurprising, since its basic meaning is a contextual tautology. Buccola points out, however, that if DIs are derived, the sentence is no longer a contextual tautology, as is evident from the paraphrase of the DIs in (41).

(40) #All of my grandparents are dead or alive.

(41) Some of my grandparents are dead and some of them are alive.

Explaining the oddness of (40) then requires explaining why distributive inferences are impossible. We indeed predict them to be impossible, because when exhaustified the existential alternative in (42) would entail the negation of (40), which would make it a contextual contradiction.[25]

(42) #Some of my grandparents are dead or alive.

The same pattern is found with modal quantification:

(43) a. #He must be dead or alive.
b. He might be dead and he might be alive.
c. #He might be dead or alive.

25 An alternative explanation for the unacceptability of (40) comes from work by Bassi et al. (2021) who argue that the contribution of $\mathcal{E}xh$ is presuppositional. If distributive inferences are presuppositions, then whenever they are satisfied (40) is a contextual tautology.

8 Concluding remarks

We have argued for an account of the Categorical Denić Generalization (CDG) which is based on the observation that when distributive inferences are derived, existential and universal alternatives become equivalent whenever $n \geq |D|$. We proposed a requirement on Sentence-Question Correspondence which rules out situations where existential and universal alternatives are equivalent. Our proposal provides a general theory which covers both the CDG and other much discussed cases where existential and universal alternatives are contextually equivalent, and also makes novel predictions which we've been able to corroborate (mostly).

Our focus in this paper has been the categorical effect which blocks distributive inferences when $n \geq |D|$. Recall, however, that Denić (2021) focused on her gradable generalization in (2), on the basis of which she argued that probabilistic considerations must enter into the computation of SIs. To conclude this paper, we would like to briefly discuss this argument, given that (as we demonstrated) the categorical effect can be accounted for with no need for probabilistic considerations. Denić's argument is important because we are not aware of any other argument that probabilities play a role in the computation of SIs; in fact, as has been pointed out in recent work, letting probabilities play a role in the computation of SIs often leads to problematic results (see especially Fox and Katzir 2021; Cremers et al. 2022).[26] A modular view where SI computation is blind to probabilities is thus motivated, but it stands in direct conflict with Denić's argument based on the gradable effect. We think, however, that the gradable effect can be explained within a modular view where probabilities only affect disambiguation. In what follows we briefly sketch how such an explanation might look.

As we have assumed throughout (with Denić 2021), sentences of the form $\forall x(Px \lor Qx)$ are ambiguous between having distributive inferences and having ignorance inferences: If distributive inferences are derived, then the speaker ends up committed to the truth of the proposition in (44a); if ignorance inferences are derived, then the speaker ends up committed to the disjunction of the (mutually incompatible) propositions in (44), and is also committed to being ignorant about which one of them is true.[27]

(44) a. $\forall x(Px \lor Qx) \land \exists x Px \land \exists x Qx \land \neg\exists x(Px \land Qx)$
b. $\forall x(Px \lor Qx) \land \forall x Px \land \neg\exists x Qx$
c. $\forall x(Px \lor Qx) \land \forall x Qx \land \neg\exists x Px$

Let us now make two rather mundane assumptions about the workings of disambiguation. First, we assume that it is affected by what the hearer thinks is likely to be the speaker's epistemic state: A hearer prefers a parse if it corresponds to an epistemic state which they think is most likely among

26 Chemla and Romoli (2015) argue that probabilities are relevant for theories of SIs, but their claim is compatible with a modular view in which probabilistic considerations only enter at the level of ambiguity resolution, as we will assume below.

27 As we did before, we ignore here the possibility of alternative pruning; nothing substantial hinges on this.

the competitors (those that correspond to other parses). Second, we assume that the more likely it is that a certain proposition is true, the more likely it is that the speaker's epistemic state would entail it. Given these assumptions, we can explain the gradable effect by utilizing the key observation from Denić's account of this effect, namely that given the truth of $\forall x \in D(P_1 x \vee \ldots \vee P_n x)$, the likelihood that the distributive inferences are true increases as $|D| - n$ does; in other words, as $|D| - n$ increases, the likelihood that the proposition in (44a) is true increases relative to the likelihood that one of the propositions in (44b)-(44b) is true. As $|D| - n$ increases, then, the likelihood that the speaker's epistemic state entails that the distributive inferences are true increases, and accordingly the hearer will tend to prefer the parse which entails them.

Of course, there are many other reasons why one epistemic state could be more likely than another. For instance, if it is extremely unlikely that the speaker is ignorant about which proposition in (44) is true, then the parse with distributive inferences would be preferred even if $|D| - n$ is very small (of course as long as it is larger than zero, otherwise the CDG would rule this parse out). Indeed, we have already seen examples which show precisely this effect, in (4a) and (37).

Appendix A Pruning

In §6 we identified the question a sentence addresses with its set of formal alternatives. However, as is known since Horn (1972), some formal alternatives of a sentence can sometimes be ignored ('pruned'), depending on the context. As we will see below, the picture we drew in this paper becomes more complicated once we allow for pruning to shrink the set of alternatives against which Question-Sentence Correspondence is evaluated.

One way to avoid these complications could be to simply assume that Sentence-Question Correspondence applies as stated in (24) (i.e., considering all formal alternatives), even when alternatives are pruned. This assumption is however empirically inadequate; as we have pointed out in §6 (fn. 17), an empirical problem arises whenever an implicature is not associated with a weak alternative because the strong alternative is taken to be irrelevant. There is a simple way to avoid this inadequacy by assuming a version of (24) which is based on the idea of partition by exhaustification from Fox (2018), as in (45).

(45) **Sentence-Question Correspondence** (partition by $\mathscr{E}xh$):

a. A sentence S is felicitous only if there is a question Q which is a good question for S.

b. Q is a good question for S only if

(i) $Q \supseteq \{\mathscr{E}xh_{Alt(S)} S' : S' \in Alt(S)\}$.

(ii) Q partitions the context set C.

(45) does not require the set of alternatives of a sentence to be a partition of some subset of the context set, but rather merely requires that it be a set which would yield such a partition

if exhaustification applied. As a result, it does not mandate applying $\mathcal{E}xh$, nor does it put any constraints on pruning if $\mathcal{E}xh$ does in fact apply.

While this looks like a reasonable choice to make, in this appendix we would like to consider the possibility that pruning does affect Sentence-Question Correspondence, and that it requires the set of alternatives which are left after pruning to be a partition of some subset of the context set. In §A.1 we will aim to show that our proposal can be maintained even if Sentence-Question Correspondence is affected by pruning, once an apparently unrelated issue with free choice inferences is taken into account. In §A.2 we will further suggest a pruning-based restatement of our proposal which aims to account for the felicity of some sentences which have contextually contradictory alternatives.

A.1 Pruning and distinctness

If Sentence-Question Correspondence is taken to be affected by pruning, then coming up with a specific theory of pruning becomes crucial. If pruning of any alternative is always possible, we can no longer rule out cases where *some* and *all* sentences are equivalent, as we will always be able to prune the equivalent alternative and avoid a violation of Sentence-Question Correspondence; and if pruning is never possible, then implicatures should always be obligatory, contrary to fact. So if Sentence-Question Correspondence applies after pruning, we need to weaken our constraint as in (46), which assumes that pruning is possible but limited to irrelevant alternatives, relying on Lewis's (1988) notion of relevance in (47) (mainly following Magri 2009a,b; Fox and Katzir 2011):

(46) **Sentence-Question Correspondence** (weakened):

a. A sentence S is felicitous only if there is a question Q which is a good question for S.

b. Q is a good question for S only if

(i) Q contains all alternatives in $Alt(S)$ *which are relevant given* Q.

(ii) Q partitions the context set C.

(iii) S is relevant given Q.

(47) **Lewis's notion of Relevance (Aboutness)**:

The set of propositions relevant given a partition Q of the context-set is the set of all unions of cells in Q ($\{\bigcup P : P \subseteq Q\}$)

Support for the assumption that (47) governs prunability comes from several sources. Magri argues (based on the data in (23)) that when two alternatives are contextually equivalent it's impossible to prune one without the other, which is a consequence of (46) when relevance is defined as in (47). In the same spirit, Fox and Katzir argue that an alternative cannot be pruned when it's in the Boolean closure of the set of alternatives which haven't been pruned, which is also a consequence of (46) when relevance is defined as in (47).

Let us show first how this proposal explains the basic case of oddness discussed by Magri in (23).

The contextual equivalence of *some* and *all* makes sure (just like in Magri 2009a,b) that the *some* alternative cannot be ignored once *all* is relevant. And as before, Q cannot contain both alternatives and still be a partition. Furthermore, (47) renders contradictions relevant (as noted by Lewis 1988), which is why $\mathcal{E}xh(\textit{some})$ (=*some* $\wedge\neg$ *all*) also cannot be ignored in this case.

The situation becomes more involved once we consider the case we are interested in, namely the disappearance of distributive inferences when $n \geq |D|$. First, since the universal and existential alternatives are not equivalent in this case before exhaustification applies, one might suspect that the existential alternative can be ignored so that it would not be able to block the derivation of distributive inferences if $n \geq |D|$. However, note that the existential alternative $\exists x(Px \vee Qx)$ is equivalent to the disjunction of the alternatives $\exists xPx$ and $\exists xQx$. As mentioned above, relevance is closed under Boolean operations (conjunction and negation; see Fox and Katzir 2011), hence it is impossible for $\exists xPx$ and $\exists xQx$ to be relevant without $\exists x(Px \vee Qx)$ being relevant as well (because $\exists x(Px \vee Qx) \Leftrightarrow \neg(\neg\exists xPx \wedge \neg\exists xQx)$). Since the relevance of $\exists xPx$ and $\exists xQx$ is a necessary ingredient for deriving distributive inferences, it is impossible to derive distributive inferences while pruning the existential alternative $\exists x(Px \vee Qx)$.

Second, in our discussion of the equivalence between (19) and (20), repeated in (48) and (49), we assumed that the strong exclusive inference in (48a) is derived.

(48) Both of Mary's friends are French or Spanish

- a. **Strong exclusive inference**: None of Mary's friends are both French and Spanish
- b. **Distributive inferences**:
  - (i) At least one of Mary's friends is French
  - (ii) At least one of Mary's friends is Spanish

(49) Some of Mary's friends are French or Spanish

- a. **Strong exclusive inference**: None of Mary's friends are both French and Spanish
- b. **Free Choice inferences**:
  - (i) At least one of Mary's friends is French
  - (ii) At least one of Mary's friends is Spanish

While there is evidence for the existence of the strong exclusive inference (see Chemla and Spector 2011), it is not very robust and is often not derived. Furthermore, the absence of distributive inferences when $n \geq |D|$ does not seem to depend in any way on deriving the strong exclusive inference. However, the equivalence between (48) and (49) crucially relies on this inference. If (48a) and (49a) are replaced with the weak exclusive inference in (50) the equivalence no longer holds.

(50) **Weak exclusive inference**: Not all of Mary's friends are both French and Spanish

We'd like to suggest that this issue has to do with a known problem located elsewhere, namely that free choice inferences are stronger than the conjunction of the existential disjunctive alternatives. This issue has been brought up by Menéndez-Benito (2010). Consider the following example (Menéndez-Benito demonstrated the problem with her Canasta scenario using free choice items; here we replicate it with Free Choice disjunction instead):[28]

(51) *Context: The rules of the game we play say that if you see two cards on the table, you can either take both or none, but you are not allowed to take just one of them. John doesn't know the rules, so Mary puts a red card and a blue card on the table and tells him:*
In this situation you can take the blue card or the red card (but if you take one of them you must take the other one as well).

The oddness of (51) is surprising since the sentence is true both on its ordinary meaning and its strengthened meaning. Specifically, the Free Choice inferences in (52) are true in the context:[29]

(52) a. You can take the blue card
b. You can take the red card

What seems to us to be the minimal assumption needed in order to capture the surprising oddness of (51) is the following:[30]

---

28 Menéndez-Benito's problem doesn't only arise with modal quantification. For instance, (i) is odd if all the boxes that contain ice cream also contain cake, and vice versa.

(i) (If you want to have something sweet,) there are boxes with cake or ice cream in the fridge.

29 One may think that the unacceptability of (51) comes from exclusion of a conjunctive alternative — resulting in the inference that you cannot take both the blue card and the red card (at the same time); this inference has however been shown to be optional (as mentioned above), as in the following example (Simons 2005):

(i) You may sing or dance
$\not\rightsquigarrow$ It's not the case that you may sing and dance (at the same time)

Moreover, one can see that exclusion of a conjunctive alternative is not what is at stake by considering minor variations on the same scenario where the rules permit the player to take both cards but also permit the player to take just one of the two cards (see fn. 30).

30 Menéndez-Benito (2010) herself assumes a stronger requirement than Distinctness according to which the FC inferences essentially are $\exists x(Px \wedge \neg Qx) \wedge \exists x(Qx \wedge \neg Px)$. Her stronger requirement however rules out the following example, which to us seems fine.

(i) *Context: The rules of the game we play say that if you see a blue and a red card on the table, your options are: (i) take both cards, (ii) take the blue but not the red card, or (iii) take no card; in other words, if you take the red card you must take the blue one as well. John doesn't know the rules, so Mary puts a red card and a blue card on the table and tells him:*
In this situation you can take the blue card or the red card, but if you take the red card you must take the blue card as well.

(53) **Distinctness**:

a. p and q (both of the form $\exists x\phi/\Diamond\phi$) are **distinctly satisfied** propositions in a world w *iff* p and q do not have identical witness sets.

b. A sentence S with Free choice/Distributive inferences is only judged true in a world w if the set of its Free Choice/Distributive inferences is pairwise distinctly satisfied.

(51) violates Distinctness: The set of worlds compatible with the rules where you take the blue card is the same as the set of worlds compatible with the rules where you take the red card.[31] Intuitively, in the case of (48) and (49), Distinctness requires that there would be one French friend (who is possibly also Spanish) and *another* Spanish friend (who is possibly also French). Under the distinctness assumption, (48) and (49) are equivalent even in the absence of the strong exclusive inference, since for (49b) to be true there would have to be one French friend and another Spanish friend. But if that's the case then (48b) is true. More generally, the following holds:

(54) In all worlds where $\{\exists x \in D(P_i x) : 1 \leq i \leq n\}$ is pairwise distinctly satisfied: (54a) and (54b) are equivalent *iff* $n \geq |D|$.

a. $\forall x \in D(P_1 x \vee \ldots \vee P_n x) \wedge \bigwedge\limits_{1\leq i\leq n} \exists x \in D(P_i x)$

b. $\exists x \in D(P_1 x \vee \ldots \vee P_n x) \wedge \bigwedge\limits_{1\leq i\leq n} \exists x \in D(P_i x)$

Given Distinctness, then, our account of the CDG extends to cases which involve pruning (whatever may be the proper explanation for Distinctness; see fn. 31 for one possibility).

A.2 Felicity with contextually contradictory alternatives

As established, we predict that having contextual contradictions as alternatives should lead to infelicity. Note furthermore that this does not change once we allow for pruning to affect Sentence-

31 Of course, one would like to know where distinctness comes from. One possibility relies on Singh (2008). Singh argues for the following constraint:

(i) **Incompatibility**: In a disjunction of the form $p \vee q$, the disjuncts must be incompatible with each other, i.e., $p \cap q = \emptyset$.

As evidence he presents the following examples:

(ii) a. #John ate some or not all of the cookies
b. #Mary lives in Russia or in Asia

Note that Incompatibility implies Distinctness:

(iii) If P and Q are incompatible, then in every world w, if $\exists x P x$ and $\exists x Q x$ are both true in w, they are distinctly satisfied in w.

As far as we can tell, Incompatibility will not immediately explain the original data Menéndez-Benito (2010) was concerned with, which does not involve disjunction but rather Free Choice Items.

Question Correspondence as in (46), since contradictions are always relevant and, as a result, cannot be pruned (Lewis 1988). As noted in fn. 18, there are however felicitous sentences which have contextual contradictions as alternatives. Consider the following example brought up by Roni Katzir, p.c.:

(55) Mary has five broken fingers on her right hand.
Contextually contradictory alternative: Mary has six broken fingers on her right hand.

We acknowledge that this might be a serious problem for the perspective we proposed in this paper; in what follows we will only attempt to suggest a tentative solution.

One difference between the contextual contradictions leading to infelicity (or to blocking distributive inferences) we discussed in this paper and (55) is that in (55) the contextual contradiction is not the result of an application of $\mathcal{E}xh$. We could propose to exempt alternatives which are contextually trivial from the get go from even being considered. This, in turn, could be achieved by replacing the requirement to have Sentence-Question Correspondence (in any of its versions) with a constraint on exhaustified meanings, as follows:

(56) **Constraint on exhaustified meanings**:
$\mathcal{E}xh_C(S)$ is only licensed if:
a. $\{\mathcal{E}xh_C(A) : A \in C\}$ partitions some subset of the context set,
b. C contains all alternatives in $Alt(S)$ which are relevant *and non-trivial*, and
c. S is relevant given Q.

This constraint requires that the pointwise application of $\mathcal{E}xh$ over the chosen set of alternatives be a partition, and that its domain must contain all relevant alternatives which are not trivial. (56) still explains the main facts we aim to capture in this paper if taken together with Magri's assumption that applying $\mathcal{E}xh$ is obligatory. At the same time, it is compatible with the felicity of (55) because it allows pruning of alternatives that are contextually contradictory before exhaustification.[32]

References

Alonso-Ovalle, L. (2005). Distributing the disjuncts over the modal space. In L. Bateman and C. Ussery (Eds.), *North East Linguistic Society (NELS)*, Volume 35, pp. 1–12.

Anvari, A. (2018). Logical integrity: from maximize presupposition! to mismatching implicatures. Manuscript.

Anvari, A. (2022). Atoms and oddness. Talk given at Rutgers University, April 1, 2022.

Bar-Lev, M. E. and D. Fox (2016). On the global calculation of embedded implicatures. Poster presented at MIT workshop on exhaustivity.

32 Note that exhaustifying over a simple disjunction of the form $p \vee q$ would require pruning the alternatives p and q, since otherwise exhaustification does not lead to a partition and (56) is violated.

Bar-Lev, M. E. and D. Fox (2020). Free choice, simplification, and innocent inclusion. *Natural Language Semantics 28*, 175–223.

Bassi, I., G. Del Pinal, and U. Sauerland (2021). Presuppositional exhaustification. *Semantics and Pragmatics 14*, 11.

Buccola, B. and A. Haida (2019). Obligatory irrelevance and the computation of ignorance inferences. *Journal of Semantics 36*(4), 583–616.

Chemla, E. (2009). Similarity: Towards a unified account of scalar implicatures, free choice permission and presupposition projection. Unpublished manuscript.

Chemla, E. and J. Romoli (2015). The role of probability in the accessibility of scalar inferences.

Chemla, E. and B. Spector (2011). Experimental evidence for embedded scalar implicatures. *Journal of semantics 28*(3), 359–400.

Chierchia, G. (2004). *Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface*, pp. 39–103. Oxford, UK.

Chierchia, G., D. Fox, and B. Spector (2012). Scalar implicature as a grammatical phenomenon. In C. Maienborn, K. von Heusinger, and P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning*, Volume 3, pp. 2297–2331. Mouton de Gruyter.

Chomsky, N. (1995). Bare phrase structure. In G. Webelhuth (Ed.), *Government and Binding Theory and the Minimalist Program*, pp. 385–439.

Cremers, A., E. G. Wilcox, and B. Spector (2022). Exhaustivity and anti-exhaustivity in the RSA framework: Testing the effect of prior beliefs. Ms., Vilnius University, Harvard University, and Institut Jean Nicod.

Crnič, L., E. Chemla, and D. Fox (2015, Dec). Scalar implicatures of embedded disjunction. *Natural Language Semantics 23*(4), 271–305.

Denić, M. (2021). Probabilistic informativeness in implicature computation: the case of embedded disjunctions. Manuscript.

Eckardt, R. (2007). Licensing or. In *Presupposition and implicature in compositional semantics*, pp. 34–70. Springer.

Fox, D. (2007). Free choice and the theory of scalar implicatures. In U. Sauerland and P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics*, pp. 71–120. London: Palgrave Macmillan UK.

Fox, D. (2016). On why ignorance might be part of literal meaning—commentary on Marie Christine Meyer. Talk given at the MIT workshop on exhaustivity.

Fox, D. (2018). Partition by exhaustification: Comments on Dayal 1996. In U. Sauerland and S. Solt (Eds.), *Proceedings of Sinn und Bedeutung 22*, Volume 1 of *ZASPiL 60*, pp. 403–434. Berlin: Leibniz-Centre General Linguistics.

Fox, D. and R. Katzir (2011, Mar). On the characterization of alternatives. *Natural Language Semantics 19*(1), 87–107.

Fox, D. and R. Katzir (2021). Notes on iterated rationality models of scalar implicatures. *Journal of Semantics 38*(4), 571–600.

Horn, L. R. (1972). *On the semantic properties of logical operators in English*. Ph. D. thesis, University of California, Los Angeles.

Kamp, H. (1974, jun). Free choice permission. *Proceedings of the Aristotelian Society 74*(1), 57–74.

Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy 30*(6), 669–690.

Katzir, R. (2014). On the roles of markedness and contradiction in the use of alternatives. In S. P. Reda (Ed.), *Pragmatics, Semantics and the Case of Scalar Implicatures*, pp. 40–71. Springer.

Katzir, R. and R. Singh (2015). Economy of structure and information: Oddness, questions, and answers. In E. Csipak and H. Zeijlstra (Eds.), *Proceedings of Sinn und Bedeutung*, Volume 19, pp. 302–319.

Klinedinst, N. W. (2007). *Plurality and possibility*. Ph. D. thesis, University of California, Los Angeles.

Kratzer, A. and J. Shimoyama (2002). Indeterminate pronouns: The view from japanese. In Y. Otsu (Ed.), *The Third Tokyo Conference on Psycholinguistics*, pp. 1–25. Hituzi Syobo.

Lewis, D. (1988). Relevant implication. *Theoria 54*(3), 161–174.

Magri, G. (2009a). *A theory of individual-level predicates based on blind mandatory implicatures: constraint promotion for optimality theory*. Ph. D. thesis, Massachusetts Institute of Technology.

Magri, G. (2009b). A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural language semantics 17*(3), 245–297.

Magri, G. (2011). Another argument for embedded scalar implicatures based on oddness in downward entailing environments. *Semantics and Pragmatics 4*(6), 1–51.

Menéndez-Benito, P. (2010). On universal free choice items. *Natural Language Semantics 18*(1), 33–64.

Meyer, M.-C. (2013). *Ignorance and grammar*. Ph. D. thesis, Massachusetts Institute of Technology.

Romoli, J. (2012). *Soft but Strong. Neg-raising, Soft triggers, and Exhaustification*. Ph. D. thesis, Harvard University.

van Rooij, R. (2002). Relevance only. In *Proceedings of Edilog*.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics 1*, 75–116.

Sauerland, U. (2004a). On embedded implicatures. *Journal of cognitive science 5*(1), 107–137.

Sauerland, U. (2004b). Scalar implicatures in complex sentences. *Linguistics and philosophy 27*(3), 367–391.

Simons, M. (2005). Dividing things up: The semantics of or and the modal/or interaction. *Natural Language Semantics 13*(3), 271–316.

Singh, R. (2008). On the interpretation of disjunction: Asymmetric, incremental, and eager for inconsistency. *Linguistics and Philosophy 31*(2), 245–260.

Singh, R. (2009). 'maximize presupposition!'and informationally encapsulated implicatures. In *Proceedings of Sinn und Bedeutung*, Volume 13, pp. 513–526.

Spector, B. (2006). *Aspects de la pragmatique des opérateurs logiques*. Ph. D. thesis, Paris 7.

Spector, B. (2014). Scalar implicatures, blindness and common knowledge: Comments on magri (2011). In *Pragmatics, Semantics and the Case of Scalar Implicatures*, pp. 146–169. Springer.

von Wright, G. H. (1968). *An essay in deontic logic and the general theory of action*. Amsterdam: North-Holland.