Semantics & Pragmatics Volume 4, Article 6: 1–51, 2011
http://dx.doi.org/10.3765/sp.4.6

Another argument for embedded scalar implicatures based on oddness in downward entailing environments*

Giorgio Magri
IJN, ENS, CNRS

Received 2011-05-06 / Decision 2011-07-06 / Revision Received 2011-10-04 / Decision 2011-10-07 / Revision Received 2011-10-11 / Published 2011-11-16

Abstract In Magri 2009a, I argue that a sentence such as #*Some Italians come from a warm country* sounds odd because it triggers the scalar implicature that not all Italians come from a warm country, which mismatches with the piece of common knowledge that all Italians come from the same country. If this proposal is on the right track, then oddness can be used as a diagnostic for scalar implicatures. In this paper, I use this diagnostic to provide one more argument that scalar implicatures are computed not only at the matrix level but also in embedded position. The argument is based on a puzzling pattern of oddness in downward entailing environments. Some apparently unrelated facts about restrictions on temporal modification with individual-level predicates are shown to fit into the pattern.

Keywords: Scalar implicatures; downward entailing environments; individual-level predicates; life-time effect

* This work owes a lot to discussions with Danny Fox, Gennaro Chierchia, and Benjamin Spector. I also would like to acknowledge feedbacks from the audiences at the following venues, where parts of this work have been presented: the MIT-Paris Workshop on Presuppositions and Implicatures (May 2007); the Language and Cognition reading group of Harvard University (October 2009); the workshop on individual-level predicates organized at the Maison Suger in Paris by Angelika Kratzer (December 2009); the IGG 36 (February 2010); and SALT 20 (May 2010). The paper has benefited from detailed comments from the editors of S&P and three anonymous reviewers. Some of the material presented in this paper has appeared in two conference proceedings (Magri 2009d and Magri 2011); the discussion of temporal modification with individual-level predicates in Section 5 appeared in my dissertation (Magri 2009b; see also Magri 2009c). This work was supported in part by a 'Euryi' grant from the European Science Foundation ("Presupposition: A Formal Pragmatic Approach" — Philippe Schlenker).

1 Introduction

The (a) sentences in (1)-(3) have an existential quantifier and sound odd (I use "#" as a diacritic for *oddness*); the variants in (b) sound impeccable. This contrast is puzzling. Take for instance the pair in (1). We know that all Italians come from the same country. Thus, (1a) and (1b) convey the same information, namely that Italy is warm. Why is it then that only (1b) sounds fine?[1]

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

(2) *Context*: In Italy, children always inherit the last name of their father.
 a. #Some of the children of that couple have a funny last name.
 b. The children of that couple have a funny last name.

(3) *Context*: Prof. Smith assigns the same grade (possibly a different one every term) to all of his students. [Emmanuel Chemla, as p.c. to Schlenker (2011)]
 a. #This year, Prof. Smith assigned an A to some of his students.
 b. This year, Prof. Smith assigned an A to his students.

1 Some of these cases of odd existential quantification remain odd if the existential quantifier *some* is replaced with an overt universal quantifier *all*/*every*. Some examples are in (i).

(i) a. #Every Italian comes from a warm country.
 b. #Every child of that couple has a funny last name.

(ii) This year, Professor Smith gave an A to all of his students.

Yet oddness triggered by overt universal quantifiers is less robust that the oddness triggered by existential quantifiers, as also noted in Singh 2009. For instance, while the overtly universally quantified variants (i) of sentences (1a) and (2a) sound odd, the overtly universally quantified variant (ii) of sentence (3a) was reported as fine in Schlenker (2011, p.c. from Emmanuel Chemla). In Magri 2009a, Sect. 3.4, I suggest that the oddness of sentences with overt universal quantifiers such as (i) is due to a very different mechanism than the one described in this paper, namely to competition with the alternative containing the definite or the bare plural, which is preferred by *Maximize Presupposition*. According to this proposal, the fact that certain sentences with overt universal quantifiers sound odd is orthogonal to the point made in this paper. In the rest of the paper, I will thus ignore the fact that certain sentences with overt universal quantifiers sound odd, thus effectively conflating definites, bare plurals and universal quantifiers. Let me finally point out that sentences (1a)-(3a) remain odd if the existential quantifier *some* is replaced with the negative quantifier *no*; for a conjecture concerning this latter case, see Magri 2009c, Sect. 2.4.

In general, an existentially quantified sentence such as (4a) triggers the *scalar implicature* that the corresponding universally quantified alternative (4b) is false, as schematized in (5).

(4) a. John did *some* of the homework.
b. John did *all* of the homework.

(5) $\text{some}_x P(x) \rightsquigarrow \neg\text{all}_x P(x)$.

In Magri 2009a, I argued that this scalar implicature is the driving force of the oddness of (1a)-(3a). In (6), I illustrate the idea informally for the case of (1a).

(6) a. Because of existential quantification, (1a) triggers the scalar implicature that not all Italians come from a warm country.
b. But common knowledge entails that, if some Italians come from a warm country, then all of them do, because they come from the same country.
c. The oddness of (1a) thus follows from the mismatch between the implicature (6a) and the common knowledge (6b).

Let me dub (6) *Hawkins' reasoning*, as it is close to a reasoning first developed by Hawkins (1991) in a very different context; see Heim 1991, Percus 2006, and Sauerland 2008 for discussion. The starting point of this paper is the conjecture that Hawkins' reasoning (6) is on the right track. Crucially, this means that oddness can be used as a diagnostic for scalar implicatures. Let me put this idea into context.

An important recent debate concerns the issue of whether scalar implicatures can be computed in embedded positions, besides the matrix position; see for instance Chierchia 2004 and Horn 2005 for an early formulation of the two competing views. This debate is complicated by methodological challenges. In fact, it is hard to obtain clear, crisp *direct* judgments concerning the availability of a certain implicature in a certain discourse context. And the challenge is even harder for the case of embedded implicatures, as they are less robust and the sentences needed to generate them are more complicated. One strategy to overcome this methodological challenge is to resort to psycholinguistic experimental techniques; see for instance Geurts & Pouscoulous 2009 and Chemla & Spector 2011. Another strategy is to resort to more *indirect* ways to probe into speakers' intuitions, by means of suitable diagnostics for the presence of a scalar implicature. Chierchia, Fox & Spector (to appear) provide an elegant example of this indirect methodology. It relies on the paradigm in (7), based on an observation by Hurford (1974). Sentence (7a) sounds odd. Plausibly, that is because the second disjunct (*John comes from Paris*) entails the first one (*John comes from France*). Thus, there has got to be a constraint that bans disjunctions with such a property. Call it *Hurford Constraint*. But then how come that (7b) sounds fine, given that again the second disjunct (*John did all*) entails the first one (*John did some*)? Here is a natural hypothesis. Suppose that the scalar implicature (5) can be derived not only at the matrix level but also embedded underneath disjunction. Because of this embedded implicature, sentence (7b) is effectively equivalent to (7c). As there is no entailment between the two disjuncts in (7b) and (7c), Hurford Constraint is not violated.

(7) a. #John comes from France or from Paris.
    b. John did some or all of the homework.
    c. John did only some or all of the homework.

The crucial idea here is that Hurford Constraint allows the implicature triggered in embedded position in (7b) to be detected *indirectly*, through a robust acceptability judgment. Under the assumption (6) that certain patterns of oddness are due to scalar implicatures, these judgments on oddness should also be able to be used as a robust diagnostic into the availability of implicatures. Building on this intuition, this paper develops a new tool to detect scalar implicatures in embedded positions.

The paper is organized as follows. To set the stage, Section 2 reviews the formalization of Hawkins' reasoning (6) that I have proposed in Magri 2009a, as well as in Magri 2009b,c. And Section 3 defends the main assumptions needed with a new argument, based on properties of overt *only*. Section 4 then introduces a new intriguing pattern of oddness in downward entailing environments. Section 5 provides more evidence in favor of the observed pattern, by looking at some apparently unrelated restrictions on temporal modification with individual-level predicates, previously discussed by Kratzer (1995), Percus (1997), Musan (1997), and Maienborn (2004) among others. Section 6 then argues that the observed pattern falls into place along the lines of my formalization of Hawkins' reasoning, once we assume that scalar implicatures are computed not only at the matrix level but also in embedded positions. In the end, the paper thus presents a new argument in favor of embedded scalar implicatures, as summarized in the concluding Section 7.

2 Sketch of a theory of oddness

Each of the three steps of Hawkins' reasoning (6) imposes a specific requirement on the underlying theory of scalar implicatures. Step (6a) requires the algorithm for the computation of scalar implicatures to be blind to common knowledge, otherwise the odd sentence (*Some Italians...*) and its alternative (*All Italians...*) would be indistinguishable by the algorithm, and their different status thus left unexplained. Step (6b) requires common knowledge to kick in eventually in order to derive the mismatch, but crucially only after the computation of the implicature has gone through. Finally, step (6c) requires the mismatching implicature to stay firmly in place against common knowledge, rather than being retracted and thus the odd sentence rescued from its oddness. In Magri 2009a,b,c, I discuss these three desiderata in more detail, and I sketch a theory of scalar implicatures that meets these desiderata and thus formalizes Hawkins' reasoning (6). The core assumptions specific to my proposal are reviewed in Subsection 2.2. They are stated against the background of the so called *grammatical approach* to scalar implicatures (see for instance Chierchia, Fox & Spector to appear), reviewed in Subsection 2.1.

2.1 Background assumptions

As recalled above, sentence (8a) with an existential quantifier triggers the scalar implicature (5) that the corresponding universally quantified alternative is false, namely that it is not the case that John completed all of the homework. Once this scalar implicature is factored in, sentence (8a) ends up equivalent to sentence (8b) with overt *only* (small capitals mark focus). Fox (2007) notes that this equivalence holds in full generality, as illustrated in (8)-(10).

(8) a. John did some of the homework.
    b. John only did SOME of the homework.

(9) a. John bought three houses.
    b. John only bought THREE houses.

(10) a. John talked to Mary or Sue.
    b. John only talked to Mary OR Sue.

The equivalence between the two sentences in each pair (8)-(10) is reminiscent of the equivalence between the two sentences (11). English has an overt distributivity operator *each*, that makes it possible in (11b) to apply the distributive predicate denoted by *tall* to the plural individual denoted by the definite subject *the kids*. Sentence (11a), without the overt operator *each*, sounds equivalent to the corresponding sentence (11b). A standard strategy to account for this equivalence is to assume that English has a phonologically covert variant of *each*, called the *distributivity operator*. The LF of sentence (11a) contains this covert operator and is thus identical to the LF of sentence (11b), with overt *each*.

(11) a. The kids are tall.
b. The kids each are tall.

Fox (2007) suggests that the equivalences in (8)-(10) should be handled in quite the same way as the equivalence in (11). He thus assumes that Natural Language has a phonologically covert variant of *only*, called the *exhaustivity operator* and notated EXH. And that the LF of sentences (8a)-(10a) can be endowed with this covert propositional operator EXH as in (12), in complete analogy with the LFs of the corresponding sentences (8b)-(10b), with overt propositional *only*. The meaning of the LF without the exhaustivity operator is called the *plain meaning* of the sentence while the meaning of the LF with the exhaustivity operator is called its *strengthened meaning*.

(12) [EXH/only φ]

Various arguments have been provided in the literature in favor of this assumption of a syntactically realized exhaustivity operator. One argument is that (12) allows for the exhaustivity operator to appear in embedded contexts, thus straightforwardly accounting for various patters of embedded implicatures; see Chierchia, Fox & Spector to appear. Another argument is that (12) allows for the exhaustivity operator to be iterated, as argued in Fox 2007 and Spector 2006, 2007.

The proper semantics of overt *only* and of the covert exhaustivity operator EXH has been a topic of intense research, at least since seminal work by Groenendijk & Stokhof (1984). Many proposals made in the literature share the structure in (13).[2]

(13) $$\text{EXH}(\varphi) = \varphi \wedge \bigwedge_{\psi \in \mathcal{E}xcl(\varphi)} \neg\psi$$

2 Here and throughout the paper, I sloppily use the same symbol φ for both an LF and its plain meaning, namely its standard truth conditions.

The exhaustivity operator takes a proposition φ, called the *prejacent*. And it does two things: it asserts the prejacent and it negates a bunch of alternatives ψ, namely all the alternatives ψ that belong to the set $\mathcal{E}xcl(\varphi)$ of *alternatives excludable w.r.t.* the prejacent φ. Each conjunct $\neg\psi$ in (13) is called a *scalar implicature* of the prejacent φ. The set $\mathcal{E}xcl(\varphi)$ of alternatives that are excludable is a subset of the set $\mathcal{A}lt(\varphi)$ of *scalar alternatives* associated with the prejacent φ. To complete the semantics (13) for the exhaustivity operator, we thus need a way to construct the set $\mathcal{A}lt(\varphi)$ of scalar alternatives and to carve the subset $\mathcal{E}xcl(\varphi)$ of those alternatives that are excludable.

Two main definitions of the set $\mathcal{A}lt(\varphi)$ of scalar alternatives of the prejacent φ have been considered in the literature. One dates back to at least Horn 1972. The idea is that scalar implicatures are triggered by designated lexical *scalar items*, such as *some*, *or*, numerals, etcetera. For each scalar item, the lexicon encodes a predefined set of *Horn-mates*. For instance *all* is a pre-assigned Horn-mate of *some*, *and* of *or*, and so on. The set $\mathcal{A}lt(\varphi)$ of scalar alternatives of the prejacent φ is then defined in terms of scalar items and Horn-mates as in (14).

(14) The set $\mathcal{A}lt(\varphi)$ of *scalar alternatives* of the prejacent LF φ consists of those LFs that can be obtained from the target LF φ by replacing one or more scalar items in φ with their Horn-mates.

A recent proposal due to Katzir (2007) does away with lexically predefined scalar items and Horn-mates. He assumes that the set $\mathcal{A}lt(\varphi)$ of scalar alternatives of the prejacent φ is the collection of all those LFs that are not more complex (in a certain technical sense) than the target LF φ. The proposal developed in this paper is compatible with both these approaches to the definition of the set of scalar alternatives. For concreteness, I will stick with the classical definition (14). My assumptions on lexical scalar items and Horn mates are fairly standard, basically just that an existential and a universal quantifier (of the same semantic type) are Horn-mates.

Out of the set of scalar alternatives $\mathcal{A}lt(\varphi)$, we need to carve the subset $\mathcal{E}xcl(\varphi)$ of alternatives that get actually excluded by the exhaustivity operator (13). Two main approaches have been pursued in the literature concerning the proper definition of the subset $\mathcal{E}xcl(\varphi)$ of excludable alternatives. One approach has its roots in the classical (neo)-Gricean approach to scalar implicatures. According to this approach, scalar implicatures arise because the speaker's utterance φ is compared with an alternative ψ that the speaker could have uttered instead and that would have made a priori a better utterance. The most plausible reason why the speaker did not assert that better alternative ψ is that it is false. Thus, the assertion of φ implies the negation of ψ. If conversation is construed as a game that maximizes information exchange between speaker and addressee, it makes sense to assume that an alternative ψ is "better" than the target φ in case ψ would have provided more information than φ, in the sense that ψ asymmetrically entails φ. Building on this tradition, it would be natural to define the set $\mathcal{E}xcl(\varphi)$ as consisting of those scalar alternatives $\psi \in \mathcal{A}lt(\varphi)$ such that ψ *asymmetrically entails* φ; see for instance Horn 1972. Yet, the hypothesis that implicatures are derived not through pragmatic, extra-grammatical reasoning, but rather through a syntactically realized covert operator leads to a different approach. From the latter perspective, the restriction to the set of excludable alternatives in (13) is just a pre-processing step to ensure that applying the exhaustivity operator will not get us into trouble. Excludable alternatives don't need to be "better" than the prejacent, as long as they are not "harmful" to it. In other words, there is no need for excludable alternatives to asymmetrically entail the prejacent φ. We just need the negation of the excludable alternatives to be consistent with the prejacent, so that the overall strengthened meaning (13) won't be a contradiction. This requires in particular condition (15) to hold.

(15) The set $\mathcal{E}xcl(\varphi)$ of alternatives *excludable w.r.t.* the prejacent φ consists of those scalar alternatives $\psi \in \mathcal{A}lt(\varphi)$ such that ψ can be negated *consistently* with φ.

Of course, the latter condition (15) is not sufficient in order to ensure that the strengthened meaning is not contradictory. In fact, suppose that the prejacent φ comes with two alternatives ψ_1 and ψ_2. That both alternatives can be individually negated consistently with the prejacent (in the sense that neither $\varphi \wedge \neg\psi_1$ nor $\varphi \wedge \neg\psi_2$ is a contradiction). But that they cannot be jointly negated consistently with the prejacent (in the sense that $\varphi \wedge \neg\psi_1 \wedge \neg\psi_2$ is a contradiction). The formulation in (15) does not tell us how to proceed in this case. Some refinements are thus in order, as suggested for instance by Fox (2007) and Spector (2006). Yet, in all the cases considered in this paper there is only one alternative at play. Thus, even the simplified formulation in (15) will do the job. The adoption of this simplified formulation ensures furthermore that my proposal is actually compatible with various different refinements of condition (15).[3]

According to the definition (13), the exhaustivity operator is a (universal) quantifier over excludable alternatives to the prejacent. Just as it is the case for any overt quantifier, also in the case of this exhaustivity operator, the domain needs to be restricted by a contextually assigned *relevance predicate* $\mathcal{R}$. I will thus write $\text{EXH}_{\mathcal{R}}$ and slightly amend the original semantics (13) as in (16): in order for a scalar alternative to be negated by the exhaustivity operator, it need not only be excludable but also be relevant. Irrelevant alternatives don't matter, and thus there is no point in excluding them; see Fox & Spector 2009 and Fox & Katzir 2011 for further elaboration on this point.

(16) $$\text{EXH}_{\mathcal{R}}(\varphi) = \varphi \wedge \bigwedge_{\psi \in \mathcal{R} \cap \mathcal{E}xcl(\varphi)} \neg\psi$$

The assumptions (12)-(16) just listed have been defended independently, as indicated by the references provided. I take them to jointly characterize the so called *grammatical approach* to scalar implicatures, as reviewed for instance in Chierchia, Fox & Spector to appear. On top of these background assumptions, I would like to add three more assumptions, specific to my proposal.

2.2 Specific assumptions

In order to compute the strengthened meaning of a prejacent φ, we need to determine for each scalar alternative ψ whether it is excludable or not. According to definition (15), that means that we have to determine whether the negation of the alternative ψ contradicts the prejacent φ. In order to do that, we have available two notions of contradictoriness, namely *logical* contradictoriness as well as contradictoriness *relative to common knowledge*,[4] as defined in (17). The two notions differ because the latter only looks at

3 The choice between the classical definition of excludable alternatives in terms of asymmetric entailment and the more recent one in terms of non-contradictoriness matters only slightly for the proposal made in this paper; see the discussion in footnote 27.

4 I don't make any distinction between *common knowledge* and *common ground*. As far as I can tell, the proposal developed in this paper is perfectly compatible with recent technical developments of the notion of common ground, such as those in Stalnaker 2002. The proposal is also compatible with the simplest possible construal of this notion, according to which common ground and common knowledge are nothing but a set of possible worlds, namely the set of possible worlds consistent with the assumptions currently made in the discourse.

those possible worlds that are consistent with all assumptions made in the discourse.

(17) a. The negation of ψ *logically* contradicts φ iff there exists no possible world where φ is true and ψ false.

b. The negation of ψ contradicts φ *given common knowledge* iff there exists no possible world compatible with common knowledge where φ is true and ψ false.

Which of these two notions of contradictoriness (17) is the one relevant for the computation (15) of excludable alternatives? As stated in (18), I submit that it is the notion of logical contradictoriness (17a), not the notion of contradictoriness (17b) sensitive to common knowledge.

(18) The computation of excludable alternatives is *blind* to common knowledge, in the sense that excludable alternatives are those alternatives ψ that are *logically* consistent with the negation of the prejacent φ.

Section 3 below will provide some evidence for this assumption (18); further independent evidence is provided by Fox & Hackl (2006).

The domain of alternatives that the exhaustivity operator (16) quantifies over is restricted by a relevance predicate $\mathcal{R}$. I take $\mathcal{R}$ to be a free variable, whose value is assigned by context. Yet, valid assignments must satisfy certain grammatical axioms. In particular, I submit the two axioms (19).

(19) a. The prejacent of the exhaustivity operator is relevant.

b. If two propositions are contextually equivalent, then they pattern alike w.r.t. relevance, namely they are both relevant or else both irrelevant.

Axiom (19a) might be related to Grice's (1975) *Maxim of Relevance*.[5] And it could be formalized as a presupposition triggered by the exhaustivity opera-

5 Let me comment on the relationship between the Maxim of Relevance and my axiom (19a). Consider for instance an utterance to the effect that John did some of the homework. The Maxim of Relevance says that this utterance better be relevant. Yet, from the perspective adopted in this paper, what this utterance consists of is not just the prejacent φ = [John did some] but rather the structure [EXH φ] with an exhaustivity operator. Thus, the Maxim of Relevance just says that the latter LF is relevant, not that the prejacent φ by itself is relevant, as demanded by my axiom (19a). It might be possible to derive the axiom (19a) from the Maxim of Relevance through some further constraint to the effect that a sentence cannot be relevant just in virtue of its implicatures (at least in plain, non ironic contexts). But I will not explore this option further.

tor.[6] Axiom (19b) follows from the intuition that relevance is a contextual notion and is thus closed w.r.t. contextual equivalence

The exhaustivity operator is covert. Any covert category raises the *recoverability problem* (see for instance Rizzi 1986): how do we recover whether a covert category is instantiated or not in a given LF? Let me suggest that there is not really any recoverability problem for the exhaustivity operator, because of (20).

(20) The exhaustivity operator is *syntactically mandatory* at matrix scope.

Throughout this Section, I only look at simple, monoclausal sentences. Thus, it is enough for the moment to assume that the exhaustivity operator is mandatory *at matrix level*, as stated in (20). The rest of the paper will argue for an extension of this assumption, from the matrix level to any embedded propositional site.

To conclude, let me make explicit the correspondence between the three assumptions (18), (19), and (20) just introduced and the three desiderata associated with the three steps of Hawkins' reasoning (6). Step (6a) requires the computation of scalar implicatures to be blind to common knowledge, in order for the existential prejacent (*Some Italians...*) and the universal alternative (*All Italians...*) to be distinguishable despite the fact that they are equivalent given common knowledge. This desideratum is met through assumption (18) that the set of excludable alternatives is computed blind to common knowledge. Yet, step (6b) requires common knowledge to eventually play a role, in order to derive the mismatch between the implicature and common knowledge. This desideratum is met through the assumption that the computation of implicatures depends on a relevance predicate $\mathcal{R}$ which is in turn sensitive to common knowledge, through the closure property (19b). Finally, step (6c) requires the mismatching implicature to be firmly locked into place in order to enforce oddness. This desideratum is met by assumption (20) that the exhaustivity operator is syntactically mandatory. Let me now show in detail how these assumptions together derive Hawkin's reasoning.

6 Say, by slightly modifying the original semantics (16) of the exhaustivity operator as in (i).

(i) $$\text{EXH}_{\mathcal{R}} = \lambda\varphi : \mathcal{R}(\varphi)\,.\,\varphi \wedge \bigwedge_{\psi \in \mathcal{R} \cap \mathcal{E}xcl(\varphi)} \neg\psi$$

This would make the exhaustivity operator more parallel to overt *only*, as both would trigger a presupposition on the prejacent: the latter presupposes that the prejacent is true; the former that it is relevant.

2.3 Formalization of Hawkins' reasoning

For concreteness, let me focus on the oddness of sentence (1a); analogous considerations hold of course for the other odd sentences considered in (1)-(3). By assumption (20) that the exhaustivity operator is mandatory at matrix scope, the LF of sentence (1a) is (21), with a matrix exhaustivity operator.

(21) [$\text{EXH}_{\mathcal{R}}$ [$_{\varphi}$ Some Italians come from a warm country]]

The matrix prejacent φ comes with a unique scalar alternative, namely ψ in (22), obtained from the prejacent φ by replacing *some* with the Horn-mate *all*. The two alternative propositions φ and ψ are equivalent given the piece of common knowledge that all Italians come from the same country.

(22) ψ = All Italians come from a warm country.

The negation of this alternative ψ is *logically* compatible with the prejacent φ, because of worlds such as (23), where the prejacent φ is true but the alternative ψ false, as some but not all Italians come from a warm country.

(23)

	Aldo	Giovanni	Giacomo
warm:	√		√
not warm:		√	

Of course, worlds such as (23) are not compatible with the piece of common knowledge that all Italians come from the same county. Yet, by assumption (18) that the computation of the set of excludable alternatives (15) is blind to common knowledge, this observation plays no role. Crazy worlds such as (23) are just as good as worlds compatible with common knowledge. Thus, the definition (16) of the exhaustivity operator yields the interpretation (24) for the LF (21).

(24)
$$\begin{aligned} [\![(21)]\!] &= \text{EXH}_{\mathcal{R}}(\varphi) \\ &= \varphi \wedge \big(\neg\psi \vee \neg\mathcal{R}(\psi)\big) \\ &= \underbrace{\text{some Italians c.f.w.c.}}_{\varphi} \text{ and } \big(\underbrace{\text{not all do}}_{\neg\psi} \text{ or } \underbrace{\text{that is not relevant}}_{\neg\mathcal{R}(\psi)}\big) \end{aligned}$$

The prejacent φ has got to be relevant, as required by axiom (19a) on the contextually supplied relevance predicate $\mathcal{R}$. As the alternative ψ in (22) is contextually equivalent to the prejacent φ, then axiom (19b) requires ψ to be relevant too. In fact, as φ and ψ make exactly the same contribution to context (namely both say that Italy is warm), how could one be relevant without the other being relevant too? Since ψ is relevant, namely $\mathcal{R}(\psi)$ is true, then (24) can be simplified as in (25).

(25) $[\![(21)]\!] = \varphi \wedge \neg\psi = \underbrace{\text{Some Italians c.f.w.c.}}_{\varphi} \text{ and } \underbrace{\text{not all do}}_{\neg\psi}$

The meaning derived in (25), that some but not all Italians come from a warm country, is a contextual contradiction, given the piece of common knowledge that all Italians come from the same country. The oddness of sentence (1a) thus follows from the fact that it unambiguously denotes a contextual contradiction.

2.4 Nothing changes for plain cases

Out of the blue, the negation of the alternative (26b) is consistent with the prejacent (26a) both logically and relative to common knowledge. Let me call this a *plain* case. To close this Section, let me make sure that the specific assumptions (18)-(20), that were needed to formalize Hawkins' reasoning, are compatible with plain cases.

(26) a. John did *some* of the homework.
b. John did *all* of the homework.

Assumption (18), that excludable alternatives are computed relative to logical contradictoriness rather than contextual contradictoriness, has no effect for plain cases. In fact, the negation of the alternative (26b) is consistent with the prejacent (26a) no matter which one of the two notions of consistency we pick. Also assumption (19b), that relevance is closed w.r.t. contextual equivalence, is moot for plain cases. In fact, the prejacent (26a) and the alternative (26b) are not contextually equivalent.

Let me now focus on the more delicate assumption (20), that the exhaustivity operator is mandatory. In the preceding Subsection, I have suggested that certain sentences sound odd because they trigger a scalar implicature that mismatches with common knowledge. As the oddness effect triggered by these sentences is rather robust, I need these mismatching implicatures to be rather robust, in particular to be automatic, not context dependent and not cancelable. Assumption (20) on the obligatoriness of the exhaustivity operator is part of the machinery needed to ensure robustness of these mismatching implicatures. Yet, implicatures in plain cases are well known to be flimsy and context dependent. For instance, sentence (26a) triggers the scalar implicature that the alternative (26b) is false (namely that John did only some of the homework) in the context of the background question (27a).

But that implicature is weaker and perhaps unavailable in the case of the question (27b).

(27) a. How much homework did John do?
b. Who did some of the homework?

Within a framework that derives implicatures through a covert exhaustivity operator, the presence or absence of an implicature could be easily taken to reflect the presence or absence of the exhaustivity operator, as stated in (28). This explanation (28) for why plain implicatures are optional and context-dependent is obviously at odds with my assumption (20) that the exhaustivity operator is mandatory.

(28) We get the implicature that (26b) is false $\iff$ EXH$_{\mathcal{R}}$ is present at LF;
we don't get the implicature that (26b) is false $\iff$ EXH$_{\mathcal{R}}$ is absent at LF.

Yet, this explanation (28) for why plain implicatures are optional based on optionality of the exhaustivity operator cannot be on the right track. In fact, if the absence of the not-all implicature in the context (27b) were due to the absence of the exhaustivity operator, then we would expect no implicature at all in this context. But that's not correct: sentence (26a) in the context (27b) does trigger an implicature, namely that no one other than John did some of the homework. In other words, the conjecture (28) would predict the scalar implicatures of a given sentence to be either altogether absent or altogether present. Instead, we find certain implicatures present in contexts where other are absent. A more fine grained strategy than (28) is thus needed.

Following van Kuppevelt (1996), van Rooij (2001), and Fox & Spector (2009) (see also Zondervan 2010 for a review), I will thus entertain the following alternative account (29) for the context sensitivity of the implicature associated with the alternative (26b): in certain contexts, the alternative (26b) is not relevant (namely it does not belong to $\mathcal{R}$) and thus does not get negated by the exhaustivity operator. From this perspective, implicature cancellation is just a special case of run of the mill contextual domain restriction.[7]

(29) We get the implicature that (26b) is false $\iff$ (26b) $\in \mathcal{R}$;
we do not get the implicature that (26b) is false $\iff$ (26b) $\notin \mathcal{R}$.

This account (29) for the flimsiness and context dependence of plain scalar implicatures is perfectly compatible with my assumption (20) that the exhaustivity operator is mandatory. And the difference between the stubbornness

7 Thanks to Irene Heim (p.c.) for helping me clarify this point.

of mismatching implicatures and the flimsiness of plain ones is readily explained. In plain cases, the scalar alternative might not be relevant and thus the implicature is optional, despite the presence of the exhaustivity operator. But in the odd cases, the alternative is necessarily relevant (because contextually equivalent to the prejacent) and thus the corresponding implicature is mandatory.[8]

8 Let me hint in this footnote at slight variant of the account just presented. It has been observed that a disjunction such as (i) triggers two types of inferences. First, it triggers the inference that the speaker does not believe that John ate the cookies (he could have eaten the cake) and does not believe that John ate the cake (he could have eaten the cookies). Second, it triggers the inference that the speaker believes that John did not eat both. The two inferences display the opposite relative scope between negation and the speaker belief operator. The former is an ignorance inference of the form "it is not the case that the speaker believes that..." (in brief: $\neg \mathsf{B}_{\text{speaker}}\psi$). The latter has instead the form "the speaker believes that it is not the case that..." (in brief: $\mathsf{B}_{\text{speaker}}\neg\psi$). Following Sauerland (2004c), I call the former a *primary implicature* and the latter a *secondary implicature*.

(i) John ate the cookies or the cake.

The account for oddness I have developed so far is stated at the level of secondary implicatures. I have suggested that sentence (iia) sounds odd because it triggers the secondary implicature (iib) and the latter secondary implicature mismatches with common knowledge. The distinction between primary and secondary implicatures suggests a natural variant of this account (suggested to me by Benjamin Spector p.c.). According to this variant, sentence (iia) would sound odd because it triggers the primary implicature (iic). This primary implicature (iic) is weaker than the secondary implicature (iib). Nonetheless, it is still enough to derive a mismatch with common knowledge. In fact, common knowledge entails that, if the speaker believes the truth of the prejacent (iia), then he also believes the truth of the universally quantified alternative (as the two are equivalent w.r.t. common knowledge), contrary to what stated by the ignorance implicature (iic).

(ii) a. #Some Italians come from a warm country.
b. The speaker believes that it is not the case that all Italians come from a warm country.
c. It is not the case that the speaker believes that all Italians come from a warm country.

In order for an account for oddness to go through, it is crucial that the mismatching implicature (be it primary or secondary) is kept firmly in place. It is not simple to get the mismatching secondary implicature (iib) locked in place, because secondary implicatures are well known to display a flimsy nature and a high degree of context dependence. Thus, I had to make specific, non standard assumptions, such as the stipulation (20) on the obligatoriness of the exhaustivity operator that derives secondary implicatures. The restatement of the account at the level of primary implicatures might have an advantage from this perspective. In fact, Sauerland (2004b) notes the contrast in (iii). Sentence (iiib) cancels only the secondary implicature of sentence (iiia), namely that the speaker believes that not all of Beethoven's

3 A new argument for blindness based on overt *only*

At the beginning of Section 2, I introduced the exhaustivity operator in analogy with the distributive operator. The idea was that both operators are covert counterparts of corresponding overt operators: *only* for the case of the exhaustivity operator; *each* for the case of the distributive operator. This analogy carries over much further. Both the exhaustivity and the distributive operator are relativized to a contextual parameter: a relevance relation $\mathcal{R}$ for the exhaustivity operator; a cover C of the domain of quantification for the distributive operator, as assumed by modern theories of plural predication, such as Schwarzschild 1996. In both cases, this contextual parameter is constrained by grammatical axioms: the relevance property needs to be closed w.r.t. contextual equivalence; covers have to sum up to the entire domain. Finally, both the exhaustivity and the distributive operator can be assumed to be mandatorily present at LF: lack of implicatures does not correspond to lack of the exhaustivity operator but to a proper choice of the relevance predicate $\mathcal{R}$ that makes the corresponding alternatives not relevant; analogously, a collective reading does not correspond to lack of the distributive operator but to a proper choice of the cover C (a *collective* cover), as argued by Schwarzschild.

From this perspective, the most surprising property of the proposal sketched in the preceding Section is assumption (18) that scalar implicatures are computed blind to common knowledge. This assumption seems particularly at odds with the classical Gricean intuition that implicatures are rooted

symphonies were played. Sentence (iiic) cancels also the primary implicature, namely that it is not the case that the speaker believes that all symphonies were played. The contrast in acceptability between (iiib) and (iiic) suggests that secondary implicatures can be canceled by the simple assertion of the opposite while primary implicatures cannot.

(iii) a. They played many of Beethoven's symphonies, . . .
b. . . . and possibly all.
c. . . . #and definitely all.

This observation seems to suggest that primary implicatures are harder to cancel than secondary implicatures, and thus feel more robust. If that is indeed the case, then primary implicatures might provide a better tool than secondary implicatures in order to account for oddness. This alternative account for oddness based on primary rather than secondary implicatures easily derives the patterns of oddness considered in this paper. Yet, it does not extend to other more complicated patterns considered in Magri 2009a, Sections 3.3.2 and 4.2. It is for this reason that I have not pursued this alternative account for oddness just sketched.

in general principles of rational communication. There really seems to be nothing rational in blindness to common knowledge. In this Section, I will present a new argument to turn this objection upside down. The idea of the argument is as follows. According to the framework endorsed in this paper, scalar implicatures are derived by appending to the LF a covert exhaustivity operator EXH. This operator is construed as a covert variant of overt *only*. Assume that the two operators differ from each other only minimally. For the sake of the argument, let me take the analogy perhaps a step too far, and assume that (13) is also the proper semantics of overt *only*, as stated in (30).[9]

(30) $\text{EXH}(\varphi) = [\![\text{only}]\!](\varphi) = \varphi \wedge \bigwedge_{\psi \in \mathcal{E}xcl(\varphi)} \neg\psi.$

By exploiting this analogy, this Section develops an argument for blindness of the exhaustivity operator by arguing that overt *only* is indeed blind to common knowledge. In the end, the theory of the exhaustivity operator developed in this paper is thus perfectly aligned with the theory of other Natural Language operators.

Let's consider sentence (31), obtained from the original odd sentence (1a) with the addition of overt *only*. This sentence sounds just as odd as the original sentence without overt *only*. Here is a very straightforward account of its oddness: because of overt *only*, this sentence says that some but not all Italians come from a warm country, which of course mismatches with the piece of common knowledge that all Italians come from the same country.

(31) #Only SOME Italians come from a warm country.

a. φ = Some Italians come from a warm country.
b. ψ = All Italian come from a warm country.

Which assumptions about overt *only* are needed in order to derive this result? Let φ be the prejacent of *only*, as in (31a); let ψ be the corresponding alternative with *some* replaced by *all*, as in (31b). By (30), the plain meaning of sentence (31) is thus $\varphi \wedge \neg\psi$, which indeed mismatches with common knowledge. Yet, if the computation of the set $\mathcal{E}xcl(\varphi)$ of excludable alternatives in the semantics (30) of overt *only* could take common knowledge into account, then the alternative ψ would not count as excludable, given

9 Here, I am ignoring the restriction to the relevance predicate $\mathcal{R}$, that does not play any role in the reasoning presented in this Section; also, I am ignoring the issue of the proper division of labor between assertion and presupposition in the semantics of *only*; see Ippolito 2008 for a review.

that it is contextually equivalent to the prejacent φ. I thus conclude that the proper computation of the set of excludable alternatives for overt *only* must be blind to common knowledge.

Yet, this conclusion is threatened by the following alternative account. Assume that the computation of the set of excludable alternatives for overt *only* is not at all blind to common knowledge. Thus, the alternative ψ in (31b) does not belong to the set of excludable alternatives of the prejacent φ in (31a). The oddness of sentence (31) is thus not due to any mismatch with common knowledge. Rather, the oddness of sentence (31) can be explained as follows. Since ψ in (31b) is not excludable, then the set of excludable alternatives is empty in the case of (31). Overt *only* is therefore vacuous. And sentence (31) is ruled out by the same general constraint that bans the vacuous occurrence of *only* in sentence (32).

(32) #Only EVERY boy arrived.

But this alternative line of explanation fails in cases with multiple alternatives, such as (33).[10] Suppose that the set of excludable alternatives of overt *only* is computed taking into account the common knowledge that John has an odd number of children. In this case, the alternative ψ that John has (at least)[11] three children is not excludable, since it is equivalent to the prejacent φ that John has (at least) two children in the context considered.

(33) John has an odd number of children...
...#He has only two$_F$.

a. φ = John has at least two.
b. ψ = John has at least three.
c. ψ' = John has at least four.

Yet, the occurrence of *only* in (33) is in no way vacuous, because it can still negate the alternative ψ' that John has (at least) four children. Thus, the hypothesis that the semantics of *only* is sensitive to common knowledge leads to the incorrect prediction that the sentence *John has only two* should be fine in the context considered, and furthermore should mean that John has exactly three children.

Based on these considerations, I conclude that the set of excludable alternatives for overt *only* is computed blind to common knowledge. By

10 Example (33) is based on an example pointed out to me by Danny Fox.
11 I take the fact that *only* can be construed with a numeral as evidence that numerals do have a weak *at least n* semantics.

virtue of the analogy (30) between overt *only* and the covert exhaustivity operator EXH, this conclusion lends support to my conjecture (18) that the computation of the strengthened meaning EXH(φ) of a sentence φ is blind to common knowledge. I started out with the intuition that this blindness hypothesis (18) is implausible, as the corresponding theory of implicatures would sound somewhat paradoxical. And I have concluded instead that this blindness hypothesis (18) cannot be false, as the sentence *John has (only) two children* would otherwise be able to mean in certain contexts that John has exactly three.

4 Some new facts about oddness in downward entailing environments

In Magri 2009a as well as in the preceding Sections, I have focused on simple monoclausal sentences. This section pursues the analysis of oddness further, moving from the unembedded cases (1)-(3) considered so far to embedded cases. For reasons that will be clear shortly (see the end of Subsection 4.3), the interesting case to look at is that of embedding in *downward entailing* (henceforth: DE) environments. These are environments that support inferences from the "superset" to the "subset". For instance, the restrictor of universal quantifiers is a DE environment, as shown by the fact that it supports the inference (34).

(34) Every $\underbrace{\textit{boy}}_{\text{superset}}$ did his homework $\longrightarrow$ Every $\underbrace{\textit{tall boy}}_{\text{subset}}$ did his homework.

Let φ and ψ be two contextually equivalent scalar alternatives such that ψ logically asymmetrically entails φ. The proposal presented in Section 2 predicts φ to sound odd and ψ to sound fine. In other words, if you have to choose between two contextually equivalent alternatives, you should pick the logically stronger one. One might then expect oddness to flip in DE environments. In other words, one might expect the pattern in (35) for any DE operator O_{DE}, given that $O_{\text{DE}}(\varphi)$ is logically stronger and $O_{\text{DE}}(\psi)$ logically weaker, although they are equivalent given common knowledge.

(35) ψ is a logically stronger but contextually equivalent scalar alternative to φ:

a. $O_{\text{DE}}(\varphi)$ should sound fine;
b. $O_{\text{DE}}(\psi)$ should sound odd.

Using the restrictor of universal quantifiers to investigate oddness in DE contexts, Subsection 4.1 presents some new data that seem to split into two different patterns w.r.t. prediction (35). Subsection 4.2 offers a characterization of the two patterns, based on the way contextual equivalence is achieved. Subsection 4.3 and Section 5 extend the empirical coverage of the generalization to other DE contexts. Section 6 will then argue that both patterns can be accounted for, if something like Hawkins' reasoning (6) applies also at the embedded level, thus providing evidence for embedded scalar implicatures.

4.1 Some data

Following for instance Sauerland (2004a), let me assume that the masculine gender feature is semantically vacuous. Thus, the universal quantifier in (36a)[12] has a larger restrictor (namely the entire set of Italians) while the universal quantifier in (36b) has a smaller restrictor (namely the subset of Italian women). In other words, sentence (36a) is logically stronger than sentence (36b), despite the fact that they are equivalent given the piece of common knowledge that all Italians come from the same country, both men and women. And indeed it is the logically weaker sentence (36b) that sounds odd, while the logically stronger sentence (36a) sounds fine.[13]

(36) a. Gli italiani vengono da un paese bellissimo.
The Italians-MASC come from a country beautiful
'Italians come from a beautiful country'

b. #Le italiane vengono da un paese bellissimo.
The Italians-FEM come from a country beautiful
'Italian women come from a beautiful country'

The contrast in (37) makes the same point. The universal quantifier in (37a) has a larger restrictor (namely the set of professors who assigned an A to at least *some* students) and the universal quantifier in (37b) has a smaller restrictor (namely the set of professors that assigned an A to *all* students). Thus, sentence (37a) is logically stronger than sentence (37b), despite the fact that they are equivalent given the piece of common knowledge that, if

12 For the case of (36), I assume that plural definites (and generics) are universal operators.

13 In (36), I have switched to Italian, which has overt gender morphology. Under the plausible assumption that masculine and feminine gender exponents are Horn-mates, (36a) is indeed a scalar alternative of (36b) according to the classical definition (14) of scalar alternatives. As pointed out to me by an anonymous reviewer, there is no need to switch from English to Italian if I adopt Katzir's (2007) definition of scalar alternatives. In fact, this alternative definition allows the LF *Italians come...* to be a scalar alternative of *Italian women come...*, as the former is not syntactically more complex (in Katzir's technical sense) than the latter.

some professors got a pay raise, then all professors did. And indeed it is the logically weaker sentence (37b) that sounds odd, while the logically stronger sentence (37a) sounds fine.

(37) Every year, the dean has to decide: if the college has made enough profit that year, he gives a pay raise to every professor who has assigned an A to at least some of his students; if there is not enough money, then no one gets a pay raise.

a. This year, every professor who assigned an A to some of his students got a pay raise.

b. #This year, every professor who assigned an A to all of his students got a pay raise.

These sentences fit into the scheme (38a), where R is the restrictor of the universal operator and P its nuclear scope. The two sentences in each pair only differ because of the restrictor of the universal quantifier. Thus, I will abbreviate (38a) as in (38b).

(38) a. for every x such that $R(x)$, it is the case that $P(x)$.

b. $\text{every}_x\, R(x)$.

The restrictors of the universal quantifiers in the two sentences (36a) and (36b) are *Strong* and *Weak* in (39) respectively, where I am assuming that masculine morphology is semantically vacuous. The restrictors for the two sentences (37a) and (37b) are *Strong* and *Weak* in (40). The names *Strong*/*Weak* reflect the fact that in both cases $\textit{Strong}(x)$ asymmetrically entails $\textit{Weak}(x)$.

(39) $\textit{Strong}(x)$ = x is an Italian *woman*.
$\textit{Weak}(x)$ = x is an Italian (*man or woman*).

(40) $\textit{Strong}(x)$ = x gave an A to *all* of his students.
$\textit{Weak}(x)$ = x gave an A to *some* of his students.

With the convention in (38) and the shorthands in (39)-(40), the pattern of oddness in this first set of data (36)-(37) can be schematized as in (41).

(41) $\surd\text{every}(\textit{Weak})$;
$\#\text{every}(\textit{Strong})$.

As expected by (35), the logically weaker sentence with the restrictor *Strong* sounds odd while the logically stronger alternative with the restrictor *Weak* sounds fine.

Interestingly, this is not the end of the story, though. Let's look at some more examples. The universal quantifier in (42a) has a (logically) larger restrictor (namely the set of fathers such that just some of their children have a funny last name) while the universal quantifier in (42b) has a (logically) smaller restrictor (namely the set of fathers such that all of their children have a funny last name). Thus, sentence (42a) is logically stronger than the alternative (42b), despite the fact that they are equivalent given the piece of common knowledge that all children of a given father share their father's last name. Surprisingly, it is the logically stronger sentence (42a) that sounds odd, while the logically weaker sentence (42b) sounds fine. The pair in (43) makes the same point.[14]

(42) *Context*: In Italy, children always inherit the last name of their father.

  a. #Every father some of whose children have a funny last name must pay a fine.

  b. Every father whose children have a funny last name must pay a fine.

(43) a. #Every student with a blue eye is German.

  b. Every student with blue eyes is German.

The contrast in (44) makes the same point. The universal quantifier in (44a) has a larger restrictor (namely the set of professors who assigned an A to at least *some* of the students) while the universal quantifier in (44b) has a smaller restrictor (namely the set of professors who assigned an A to *all* students). Thus, sentence (44a) is logically stronger than sentence (44b), despite the fact that they are equivalent given the piece of common knowledge that every professor assigns the same grade to all of his students. Surprisingly, it is the logically stronger sentence (44a) that sounds odd, while the logically weaker sentence (44b) sounds fine. Note that sentences (44) are identical to those in (37), the only difference being the background knowledge.

(44) *Context*: In this department, every professor assigns the same grade to all of his students.

14 Strictly speaking, example (43) makes the same point only under the assumption that the restrictor x *has blue eyes* means "x has two or more blue eyes" and is thus stronger than x *has a blue eye*. In other words, it is crucial that the plurality inference triggered by plural morphology is already in place at the relevant level of computation of oddness. This might admittedly turn out to be a tricky issue; see for instance Sauerland 2003 and Spector 2007 for discussion.

a. #This year, every professor of this department who assigned an A to *some* of his students got a prize from the dean.

b. This year, every professor of this department who assigned an A to *all* of his students got a prize from the dean.

The restrictors of the universal quantifiers in the two sentences (42a) and (42b) are *Strong* and *Weak* in (45), respectively. The restrictors for the two sentences (44a) and (44b) are *Strong* and *Weak* in (40), repeated in (46). Again, $Strong(x)$ (logically) asymmetrically entails $Weak(x)$ in both cases.

(45) $Strong(x)$ = x is a father *all* of whose children ...
$Weak(x)$ = x is a father *some* of whose children ...

(46) $Strong(x)$ = x gave an A to *all* of his students.
$Weak(x)$ = x gave an A to *some* of his students.

With the convention in (38) and the shorthands in (45)-(46), the pattern of oddness (42)-(44) can be schematized as in (47). Contrary to the initial expectation (35), the logically weaker sentence with the restrictor *Strong* sounds fine while the logically stronger alternative with the restrictor *Weak* sounds odd.

(47) #every(*Weak*);
$\surd$every(*Strong*).

Quite surprisingly, oddness in DE contexts seems to behave according to the two patterns (41) and (47) that are the opposite of each other. The contrast between (37) and (44) is particularly puzzling: the same two contextually equivalent sentences give rise to the opposite patterns of oddness (41) and (47), depending on the common knowledge considered.

4.2 A generalization about oddness in DE contexts

What is the relevant difference between the two opposite patterns (41) and (47) of oddness in DE contexts? I submit that the relevant difference is (48). To start, let's look at some of the restrictors used to obtain pattern (41). The restrictors *Weak* and *Strong* in (39) are not contextually equivalent because of the existence of Italian men; and the restrictors *Weak* and *Strong* in (40) are not contextually equivalent because the context set up in (37) does not in any way entail that professors that gave an A to some of their students also gave an A to all of them. The situation is very different if we look at the restrictors used to obtain pattern (47). The restrictors *Weak* and *Strong* in (45) are contextually equivalent because there is no father whose children have different last names; and the restrictors *Weak* and *Strong* in (46) are contextually equivalent because the context set up in (44) ensures that professors give the same grade to all of their students.

(48) a. In the case of pattern (41), the two restrictors *Weak* and *Strong* are *not* contextually equivalent (although the matrix sentences are);

b. in the case of pattern (47), the two restrictors *Weak* and *Strong* are *indeed* contextually equivalent (whereby the matrix sentences are too).

Here is another way to put it. In both patterns, the two matrix sentences are contextually equivalent. But the two patterns differ w.r.t. the level at which contextual equivalence is established: in the case of (47), it is established at the embedded level, and then projected up; in the case of (41) it is only established at the matrix level.

Putting together (41), (47) and (48), we get the generalization (49). Here, $\leftrightarrow_{ck}$ stands for equivalence given common knowledge (i.e., the two sentences carve the same subset out of the set of worlds compatible with common knowledge). The case on the right hand column agrees with the initial expectation (35), that the logically *stronger* sentence *every*(*Weak*) is fine while the logically *weaker* sentence *every*(*Strong*) is odd. As it was illustrated here with sentences (36), I will refer to this case as the *Italian Women* (henceforth: IW) pattern. The case on the left hand side column is the reverse, and does not agree with the initial expectation (35). As it was illustrated here with sentences (42), I will refer to this case as the *Italian Fathers* (henceforth: IF) pattern.

(49) $\varphi \leftrightarrow_{ck} \psi$

	Strong $\leftrightarrow_{ck}$ *Weak*	*Strong* $\not\leftrightarrow_{ck}$ *Weak*
φ = every(*Strong*)	(a) fine	(c) odd
ψ = every(*Weak*)	(b) odd	(d) fine

Is there a way to extend Hawkins' reasoning (6) to account for the whole generalization (49)? I will take up this issue in Section 6. Until then, let me collect some more evidence in favor of this empirical generalization.

4.3 Extensions

So far, I have used the restrictor of universal quantifiers as the relevant DE context. Other DE contexts lead to analogous results. For instance, nothing changes when *every* is replaced by *no*. To illustrate, I provide in (50) and (51) the variants of (37) and (44), with *every* replaced by *no* (and furthermore *some* replaced by disjunction to circumvent the problem of PPI-ness of the former).

(50) *Context*: Every year, the dean has to decide: if the college has made enough profit that year, he gives a pay raise to every professor who has taught a graduate or an undergraduate class; if there is not enough money, then no one gets a pay raise.

a. This year, no professor who taught a graduate *or* an undergraduate class got a pay raise.

b. #This year, no professor who taught a graduate *and* an undergraduate class got a pay raise.

(51) *Context*: In this department, every professor teaches both a graduate and an undergraduate class in the same field of linguistics.

a. #This year, no professor who taught graduate *or* undergraduate Semantics got a pay raise

b. This year, no professor who taught graduate and undergraduate Semantics got a pay raise.

I thus extend the initial generalization (49) from the restrictor of a universal quantifier to an arbitrary DE operator O_{DE} as in (52).

(52) $\varphi \leftrightarrow_{\mathrm{ck}} \psi$

	$Strong \leftrightarrow_{\mathrm{ck}} Weak$	$Strong \not\leftrightarrow_{\mathrm{ck}} Weak$
$\varphi = O_{\mathrm{DE}}(Strong)$	(a) fine	(c) odd
$\psi = O_{\mathrm{DE}}(Weak)$	(b) odd	(d) fine

A remark on the special case where the DE operator O_{DE} is negation is in order here. A few authors have observed that odd sentences remain odd when embedded under negation. For instance Spector (2007) construes the plurality inference triggered by plural morphology as a scalar implicature. The oddness of sentence (53a) is thus due to the fact that this implicature mismatches with the piece of common knowledge that people can marry only one person at the time. Spector notes that embedding under negation does not affect the oddness of (53a), as shown in (53b).

(53) a. #Last summer, Mario married (some) Italian girls.
  b. #Despite his family's pressure, Mario didn't marry Italian girls.

But if the generalization (52) is on the right track, then the case where the DE operator O_{DE} is negation is really not the best case to investigate the behavior of oddness in DE contexts. In fact, negation has the peculiar property that the matrix sentences $\varphi = \neg \mathit{Strong}$ and $\psi = \neg \mathit{Weak}$ are contextually equivalent iff the corresponding embedded sentences *Strong* and *Weak* are contextually equivalent too. Thus, negation does not allow us to pull apart the two cases considered in the two columns of (52).

So far, I have considered the case of embedding under a DE operator O_{DE}. The generalization (52) can of course be extended to upward entailing (henceforth: UE) operators, as stated in the revised generalization (54).[15]

(54) $\varphi \leftrightarrow_{\text{ck}} \psi$

	$\mathit{Strong} \leftrightarrow_{\text{ck}} \mathit{Weak}$	$\mathit{Strong} \not\leftrightarrow_{\text{ck}} \mathit{Weak}$	
		O is DE	O is UE
$\varphi = O(\mathit{Strong})$	(a) fine	(c) odd	(e) fine
$\psi = O(\mathit{Weak})$	(b) odd	(d) fine	(f) odd

The idea behind this generalization (54) can be informally spelled out as follows. We have to choose between these two alternatives φ and ψ, obtained by embedding under an operator O either *Strong* or *Weak*. We start from the embedded context and move up. If the two embedded constituents are *indeed* contextually equivalent, then the choice is made at the embedded level: we pick the sentence with the logically stronger embedded constituent *Strong*, no matter the monotonicity of the embedding operator O. If the two embedded constituents are *not* contextually equivalent, then we cannot make the choice at the embedded level and need instead to look one level up. In this case, the monotonicity of the embedding operator O does of course matter. In fact, we pick the alternative that is globally logically stronger. In other words,

15 Thanks to Emmanuel Chemla for discussion on this point.

I assume that also in the case where the embedding operator *O* is UE, the decision is made at the embedded level whenever the embedded constituents are contextually equivalent. Of course, in the case where the embedding operator *O* is UE, it does not make any difference whether the choice is made at the matrix or at the embedded level, because the two options lead to the same conclusion, as the matrix logically stronger alternative is the one with the logically stronger embedded alternative.

5 Further evidence based on temporal modification with ILPs

As it has been noted by many authors, temporal modification with individual-level predicates (henceforth: ILPs) is heavily restricted, as illustrated in (55a). Furthermore, Kratzer (1995) notes that temporal modification through past-tense morphology in (55b) yields the inference that John is dead. Musan (1997) dubs this inference the *life-time effect*.

(55) a. #John is tall after dinner.
b. John was tall. ⇝ John is dead

Maienborn (2004) suggests a pragmatic account for the oddness of sentences like (55a);[16] and Musan (1997) suggests a pragmatic account for the life-time effect displayed by sentences such as (55b); see also Percus 1997. In Subsection 5.1, I explore in some detail the idea of a pragmatic account for restrictions (55) on tense modification with ILPs and then argue in Subsection 5.2 that these restrictions provide further evidence for my new generalization on oddness in DE contexts.

5.1 The puzzle

Let me start with a discussion of the oddness of sentence (55a) with the tense modifier *after dinner*, repeated it in (56a) together with the fine variant (56b) without the tense modifier.

16 Maienborn actually considers sentences containing locatives rather than tense modifiers, such as (ia).

(i) a. Mary is blond in her car.
b. Mary is blond when she is in her car.

But she then notes that these locatives are always given a temporal interpretation, along the lines of the paraphrase in (ib).

(56) a. John is tall after dinner.
b. John is tall.

Following Chierchia (1995) and Magri (2009a), assume that an ILP such as *tall* comes with a time argument, just as any other predicate. Assume furthermore that this time argument gets bound by a covert *generic operator* GEN. Simplifying somewhat the complicated issue of the proper semantics of the generic operator, I will take it to have universal force.[17] The plain meaning of sentence (56a) thus boils down to φ in (57a). The generic operator GEN_t binds the time argument t of the ILP *tall*. The *nuclear scope* of the operator consists of the set of times at which John is tall. Its *restrictive scope* consists of the set of times that satisfy the temporal modifier *after dinner* at which John is located.

(57) a. $\varphi = \text{GEN}_t[\textbf{in}(j,t) \wedge [\![\text{after dinner}]\!](t)][[\![\text{tall}]\!](j,t)]$
b. $\psi = \text{GEN}_t[\textbf{in}(j,t)][[\![\text{tall}]\!](j,t)]$

Assume that sentence (56a) counts among its scalar alternatives the sentence (56b) obtained from the former by dropping the restrictor *after dinner*. This assumption fits well with recent theories of scalar alternatives (such as the one developed in Katzir 2007), according to which scalar alternatives are obtained by pruning the target LF. The plain meaning of this alternative (56b) is ψ in (57b). In conclusion, the odd sentence (56a) ends up with the strengthened meaning $\varphi \wedge \neg\psi$, which says that John is tall at after dinner times but not at other times, which clearly mismatches with the piece of common knowledge that tallness is a permanent property. This I take to be the crucial insight of Maienborn's (2004) account, restated within the framework for scalar implicatures adopted here.[18]

Musan (1997) focuses on the life-time effect triggered by ILPs in the past tense, as in sentence (55b), repeated in (58a). She suggests the implicature-

17 I take the generic operator GEN to be the covert variant of the overt quantificational adverbs *usually* or *generally*. The exact quantificational force of these operators is not crucial for my purposes. What is crucial in order to bring the data on ILPs in (55) to bear on the generalization (49) on oddness in DE environments is just that the restrictor of these operators licenses DE inferences. And that is indeed the case, as shown by the fact that NPIs can occur in the restrictor of these operators, as shown in (i).

(i) a. Students that do *any* of the homework, (usually) pass the class.
b. If John gets *any* sleep, he is usually in a good mood.

18 Her original account is framed within a very different theory of scalar implicatures, developed in Blutner 1998, 2000.

based account (59) for this effect, based on the pragmatic comparison with the corresponding variant (58b) with past tense replaced by present tense.

(58) a. Gregory was from America.
     b. Gregory is from America.

(59) a. "The speaker has expressed the proposition [(58a)]" and "the speaker is maximally informed about Gregory's being from America — in particular about the duration of Gregory's being from America."
     b. "If the speaker thought that Gregory's being from America is not over, he would have expressed the proposition [(58b)], since that would have been a more informative alternative utterance about the duration of Gregory's being from America."
     c. "Thus, the speaker couldn't have been maximally informative about Gregory's being from America unless he though that Gregory's being from America is over."
     d. "Thus, the speaker has implicated that Gregory's being from America is over. Since being from America is a property that, if it holds of an individual at all, holds of that individual over its entire lifespan, and since the speaker has implicated that Gregory's being from America is over, the speaker has implicated furthermore that Gregory is dead."

The crucial step of this reasoning is (59b), which says that the alternative (58b) with the present tense is "more informative" than the alternative (58a) with the past tense. But I do not understand the way Musan argues for this claim.[19] She posits the plain meaning φ in (60a) for sentence (58a) with past tense and the plain meaning ψ in (60b) for the alternative (58b) with the past tense replaced by the present tense. Here, t is a time interval existentially quantified over.

19 Musan explains this claim in the following passage, that I do not understand: "Suppose [(58a)] is true. In this case we know the following: if [(58b)] is also true, then the situation time of *be from America* obviously reaches into the past (because of the truth of [(58a)]), i.e., the implication from the present tense clause to the past tense clause is guaranteed. But how about the case where [(58a)] is false? For practical purposes in a concrete discourse, this possibility can be disregarded because conversation takes place under the assumption that utterances are truthful. Hence, when a past tense clause is uttered, for practical purposes — which only care about cases where the past tense clause is true — the present tense clause is justified to count as more informative than the past tense clause. It seems that this relationship justifies treating past tense clauses and present tense clauses as ordered with regard to informativeness" (pp. 280-281).

(60) a. $\varphi = \exists t[t < \text{NOW} \wedge \text{AMERICAN}(g,t)]$
b. $\psi = \exists t[\text{NOW} \in t \wedge \text{AMERICAN}(g,t)]$

But why should truth-conditions ψ corresponding to present tense be more informative than truth conditions φ corresponding to past tense? The problem here is that Musan assumes the tense semantics in (61), whereby there is just no subset relationship between the two sets of times $[\![\text{PAST}]\!]$ and $[\![\text{PRES}]\!]$, and thus no way to compare the informativeness of the two corresponding propositions (58a) and (58b) w.r.t. entailment. In order to get a subset relationship, I follow Sauerland (2002) and assume that the present tense is vacuous, namely that it does not impose any restriction. This way, we do get a subset relationship $[\![\text{PAST}]\!] \subseteq [\![\text{PRES}]\!]$.

(61) a. $[\![\text{PRES}]\!] = \lambda t\,.\,\text{NOW} \in t$
b. $[\![\text{PAST}]\!] = \lambda t\,.\,t < \text{NOW}$

Yet, if we stick with Musan's assumption that the time argument of the ILP *French* is existentially quantified as in (60), then we predict sentence (58a) with past tense morphology to have stronger truth conditions and thus to be more informative than sentence (58b) with present tense, contrary to what we want. The solution to this problem consists of sticking to my initial assumption that the time argument of the ILP *French* in the two sentences (58) is bound not by an existential operator but rather by a covert generic operator with universal force. I assume furthermore that tense morphemes end up in the restrictor of this generic operator, thus deriving the two new truth conditions φ and ψ in (62) for the two sentences (58a) and (58b) with past and present tense, respectively. Now, the truth conditions φ of the sentence with past tense morphology are weaker than the truth conditions ψ of the alternative sentence with (semantically vacuous) present tense, as desired.

(62) a. $\varphi = \text{GEN}_t[\mathbf{in}(g,t) \wedge [\![\text{past}]\!](t)][[\![\text{tall}]\!](g,t)]$
b. $\psi = \text{GEN}_t[\mathbf{in}(g,t)][[\![\text{tall}]\!](g,t)]$

Is this enough to derive the life time effect? Not at all. We now predict sentence (58a) with the past tense to end up with the strengthened meaning $\varphi \wedge \neg\psi$, which says that John was tall at times in the past but he is not tall at times that are not in the past. This strengthened meaning contradicts the piece of common knowledge that tallness is a permanent property and furthermore entails that Gregory is still alive in the present. This result is the opposite of what we want: rather than predicting the sentence to sound fine and to trigger the inference that Gregory is dead, we predict it to sound odd and to trigger the inference that he is still alive. The nature of the problem pops up very clearly if we compare the two pairs of truth conditions φ/ψ in (57) and (62): they are formally identical, the only irrelevant difference being that in the former case the tense restrictor is *after dinner* while in the latter case it is PAST. How can we then make sense of the different effects triggered by the two sentences (55), namely oddness in one case and the life-time inference in the other case?

5.2 A solution based on generalization (49)

I will now argue that the empirical generalization (49) sheds some light on the puzzle of temporal modification with ILPs just presented: the oddness effect with the tense modifier *after dinner* in (55a) and the life-time effect with past tense morphology in (55b) correspond to the two columns of the generalization. To start, consider again the odd sentence (55a) with the tense modifier *after dinner* and its fine variant without it, repeated in (63). Assume that the truth conditions of these sentences are indeed (57), with a generic operator that plays the role of a universal quantifier over time. These truth conditions can be schematized as in (64), using the convention introduced in (38) as well as the shorthands defined in (65) for the two restrictors *Strong* and *Weak*. Note that this is a case where the two restrictors cannot be equivalent given common knowledge: in fact, the equivalence *Strong* $\leftrightarrow_{\text{ck}}$ *Weak* would mean that it follows from common knowledge that John is only alive at after dinner times, which obviously cannot be. Hence, it is the right hand side column of generalization (49) that applies in this case: sentence (63a), with the tense modifier and thus the logically weaker truth conditions φ with the restrictor *Strong*, is correctly predicted to sound odd; the alternative (63b), with the logically stronger truth conditions ψ with the restrictor *Weak*, is predicted to sound fine.

(63) a. φ = John is tall after dinner.
 b. ψ = John is tall.

(64) a. $\varphi = \text{every}(\mathit{Strong})$
 b. $\psi = \text{every}(\mathit{Weak})$

(65) a. $\mathit{Strong} = \lambda t\,.\,\mathbf{in}(j,t) \wedge [\![\text{after-dinner}]\!](t)$.
 b. $\mathit{Weak} = \lambda t\,.\,\mathbf{in}(j,t)$.

Let's now turn to the case of sentence (55b) that triggers the life-time effect because of the past tense, repeated in (66a), together with the corresponding present tense alternative (66b). Under the assumption (62) that the truth conditions of these sentences indeed contain a generic operator and that the generic operator plays the role of a universal quantifier over time, the truth conditions of the two sentences (66) can be schematized as in (67), using the convention introduced in (38) as well as the shorthands defined in (68) for the two restrictors *Strong* and *Weak*.

(66) a. φ = Gregory was from America.

b. ψ = Gregory is from America.

(67) a. $\varphi = \text{every}(\mathit{Strong})$

b. $\psi = \text{every}(\mathit{Weak})$

(68) a. $\mathit{Strong} = \lambda t\,.\,\mathbf{in}(g,t) \wedge [\![\textsc{past}]\!](t)$.

b. $\mathit{Weak} = \lambda t\,.\,\mathbf{in}(g,t)$.

Two cases now need to be considered. One case is that it follows from common knowledge that Gregory is already dead. This means that the two restrictors *Strong* and *Weak* in (68) are equivalent given common knowledge.[20] In this case, it is the left hand side column of generalization (49) that applies: sentence (66a) with the ILP in the past tense and thus the logically stronger restrictor *Strong* is correctly predicted to sound fine; while the alternative sentence (66b) with the ILP in the present tense and thus the logically weaker restrictor *Weak* is correctly predicted to sound odd. The other case that needs to be considered is that it does not follow from common knowledge that Gregory is already dead. This means that the two restrictors *Strong* and *Weak* in (68) are *not* equivalent given common knowledge. In this case, it is the right hand side column of generalization (49) that applies: sentence (66a) with the ILP in the past tense and thus the logically weaker truth conditions φ with the restrictor *Strong* is correctly predicted to sound odd; while the alternative sentence (66b) with the ILP in the present tense and thus the logically stronger truth conditions ψ with the restrictor *Strong* is correctly predicted to sound fine. In conclusion, sentence (66a) with the ILP in the past tense can only be used when Gregory is known to be dead, thus deriving the life-time effect.[21] The account for the generalization (49) developed in the

20 Indeed, to say that Gregory is dead means that his lifespan $\lambda t\,.\,\mathbf{in}(g,t)$ is a subset of $[\![\textsc{past}]\!]$, so that *Strong* and *Weak* in (68) are equivalent.

21 Consider the sentence *Gregory is tall*, with the present tense morphology. My proposal does not predict that common knowledge should entail that Gregory is alive, in order for

next Section thus extends to the case of the two sentences (55), completing the account for restrictions on tense modification with ILPs.

6 An account for oddness in DE contexts based on embedded implicatures

In this Section, I focus for concreteness on the initial formulation (49) of my generalization on oddness in embedded contexts, tailored to the case of the restrictor of universal quantifiers. In Subsection 6.1, I spell out in detail the challenges raised by this generalization. And in Subsections 6.2-6.4, I argue that these challenges can be met by running Hawkins' reasoning both at the matrix and at the embedded level. As Hawkins' reasoning relies on scalar implicatures, the proposal thus bears on the existence of embedded scalar implicatures. The details of the account are collected in a final Appendix.

6.1 The two challenges raised by generalization (49)

To get started, it is useful to understand in detail how Hawkins' reasoning (6) sketched in Section 2, as it stands, fails to account for generalization (49). At a close look, it turns out that this generalization (49) raises two different challenges for Hawkins' reasoning.

6.1.1 First challenge

Consider again the odd sentence (69a) of the IW pattern together with the fine sentence (69b) of the IF pattern.

(69) a. #$\underbrace{\text{Italian women}}_{\textit{Strong}}$ come from a warm country.

b. √Every $\underbrace{\text{father whose children have a funny last name}}_{\textit{Strong}}$ must pay a fine.

this sentence to be felicitously uttered. Rather, it predicts something weaker: that common knowledge should not entail that he is dead. I think that this weaker prediction might indeed be on the right track, as shown by the fact that the following dialogue seems fine.

(i) A: Do you know Gregory? do you know if he still alive?
B: That I don't know. I know he is tall and has blue eyes.

As shown in (70), Hawkins' reasoning straightforwardly derives a mismatching implicature that accounts for the oddness of sentence (69a) of the IW pattern. By (20), the LF comes with a matrix exhaustivity operator, as in (70a). By (19), the implicature is mandatory because of the contextual equivalence between the alternative and the prejacent, as in (70b). In the end, the sentence thus means a contextual contradiction, as stated in (70c).

(70) ⟦#Italian women come from a warm country⟧

$$\overset{(a)}{=} \textsc{exh}_{\mathcal{R}}\big(\underbrace{\text{All Italian women c.f.w.c.}}_{\text{every}(\textit{Strong})}\big)$$

$$\overset{(b)}{=} \underbrace{\text{All Italian women c.f.w.c.}}_{\text{every}(\textit{Strong})} \text{ and not } \underbrace{\text{all Italians c.f.w.c.}}_{\text{every}(\textit{Weak})}$$

$$\overset{(c)}{=} \text{contextual contradiction}$$

Yet, by parity of reasoning, Hawkins' reasoning unfortunately also derives a mismatching implicature for the fine sentence (69b) of the IF pattern, as shown by the analogous reasoning in (71).

(71) ⟦ √Every father all of whose children… ⟧

$$= \textsc{exh}_{\mathcal{R}}\big(\underbrace{\text{Every father all of whose children...}}_{\text{every}(\textit{Strong})}\big)$$

$$= \underbrace{\text{Every father all of whose...}}_{\text{every}(\textit{Strong})} \text{ and not } \underbrace{\text{every father some of whose...}}_{\text{every}(\textit{Weak})}$$

$$= \text{contextual contradiction}$$

Thus, the first challenge raised by generalization (49) for Hawkins' reasoning is as follows: the fine sentence (69b) of the IF pattern needs to be somehow protected from that same mismatching implicature that is used to explain the oddness of sentence (69a) of the IW pattern.

6.1.2 Second challenge

Consider again the good sentence (72a) of the IW pattern together with the odd sentence (72b) of the IF pattern.

(72) a. √$\underbrace{\text{Italians}}_{\textit{Strong}}$ come from a warm country.

b. #Every $\underbrace{\text{father some of whose kids have a funny last name}}_{\textit{Weak}}$ will be fined.

As shown in (73), Hawkins' reasoning is perfectly compatible with the fine status of sentence (72a) of the IW pattern because there is no mismatching implicature. By (20), again the LF comes with a matrix exhaustivity operator, as in (73a). But no implicature arises because the prejacent is logically strongest, as in (73b).

(73) $[\![\sqrt{\text{Italians come from a warm country}}]\!] \overset{(a)}{=} \textsc{exh}_{\mathcal{R}}(\underbrace{\text{All Italians c.f.w.c.}}_{\text{every}(\textit{Weak})}) \overset{(b)}{=} \text{all Italians c.f.w.c.}$

Yet, by parity of reasoning, Hawkins' reasoning unfortunately derives no mismatching implicature also for the odd sentence (69b) of the IF pattern, as shown by the analogous reasoning in (74).

(74) $[\![\sqrt{\text{Every father some of whose children...}}]\!] =$
$= \textsc{exh}_{\mathcal{R}}(\underbrace{\text{Every father some of whose children...}}_{\text{every}(\textit{Weak})})$
$= \text{Every father some of whose children...}$

Thus, the second challenge raised by generalization (49) for Hawkins' reasoning is as follows: a mismatching implicature needs to be cooked up for the odd sentence (72b) of the IF pattern, but in such a way that it will not affect the fine sentence (72a) of the IW pattern.

6.2 Mandatory embedded exhaustivity operators

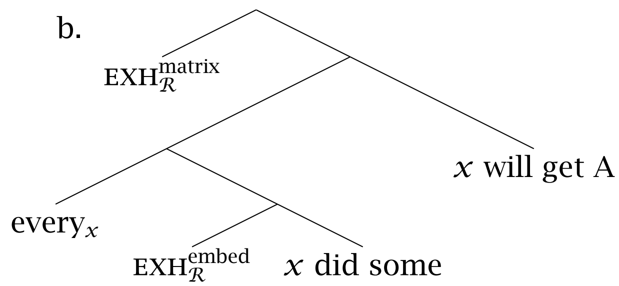
I assumed in (20) that the exhaustivity operator must mandatorily appear at LF. For the unembedded cases (1a)-(3a), it was enough that the exhaustivity operator be mandatory at matrix scope. But what is special about matrix scope? Nothing. Thus, I refine my initial assumption (20) as in (75).

(75) The exhaustivity operator $\textsc{exh}_{\mathcal{R}}$ is mandatory *at every scope site.*

Assumption (75) says, in particular, that there are embedded exhaustivity operators in DE environments. Thus, the proper LF of a sentence such as (76a) is indeed (76b), with a matrix exhaustivity operator and another one embedded in the restrictor of the universal quantifier.[22, 23]

(76) a. Everyone who did some of the homework will get an A.

b.



Assumption (75) is my only refinement of the theory sketched in Section 2. The other assumptions stay the same: the basic semantics (15)-(16) of the exhaustivity operator; the blindness assumption (18); the axioms (19) on relevance.

6.3 Sketch of the account

In Subsection 6.1, I have presented the two challenges raised by generalization (49) against Hawkins' reasoning. Both challenges were due to the fact that Hawkins' reasoning was being applied only at the matrix level. In fact, it turns out that both challenges are solved by assuming (75): that the exhaustivity operator is mandatory at every scope site and that Hawkins' reasoning therefore also applies in embedded positions. I prove this claim in detail in the Appendix. Here, let me present the idea informally.

Consider first the IW pattern, repeated in (77). This is the pattern that arises when the two restrictors *Strong* and *Weak* are *not* contextually equivalent. Sentence (77b) with a logically stronger global plain meaning sounds

22 In (16), I have construed the exhaustivity operator as a propositional operator. In order to put an exhaustivity operator in the restrictor of *every* in (76), I thus need a constituent of type t. As long as the relevant scalar item sits inside a relative clause embedded in the restrictor, as in (76a), the relative clause provides the proper argument for the exhaustivity operator. In other words, I am assuming that the wh-phrase *who* (interpreted as a λ-abstractor) moves leaving in place a trace and that the exhaustivity operator sits in between the wh-phrase and its trace. In the paper, I notate this LF compactly as in (76b), by letting the quantifier bind a variable x that saturates its restrictor. Alternatively, I could define the exhaustivity operator for arbitrary types "that end in t" and apply it above the λ-abstracting index, to the constituent of type $\langle e, t\rangle$ obtained through λ-abstraction.

23 Of course, assumption (75) also forces an exhaustivity operator in the nuclear scope of the universal quantifier. But I ignore that third instance of the exhaustivity operator here, as the nuclear scope contains no scalar item.

fine while sentence (77a) with a logically weaker global plain meaning sounds odd.

(77) a. #$\underbrace{\text{Italian women}}_{\textit{Strong}}$ come from a warm country. #every(*Strong*)

b. √$\underbrace{\text{Italians}}_{\textit{Weak}}$ come from a warm country. √every(*Weak*)

The idea of the account is as follows. Recall that what locks an implicature in place is the fact that the corresponding alternative is contextually equivalent to the prejacent. Since *Weak* and *Strong* are *not* contextually equivalent in this case, nothing forces the embedded implicature in place. Thus, the exhaustivity operator embedded in the restrictor of the universal quantifier does nothing. And the pattern of oddness in the case where *Weak* and *Strong* are *not* contextually equivalent is determined by the matrix exhaustivity operator. In conclusion, the sentence *every*(*Strong*) in (77a) with the logically weaker global plain meaning is correctly predicted to sound odd while the sentence *every*(*Weak*) in (77b) with the logically stronger global plain meaning is predicted to sound fine.

Consider next the more delicate IF pattern, repeated in (78). This is the pattern that arises when the two restrictors *Strong* and *Weak* are indeed contextually equivalent. The globally logically stronger sentence (78b) sounds odd and the globally logically weaker sentence (78a) sounds fine.

(78) a. √Every $\underbrace{\text{father all of whose children...}}_{\textit{Strong}}$ will pay a fine. √every(*Strong*)

b. #Every $\underbrace{\text{father some of whose children...}}_{\textit{Weak}}$ will pay a fine. #every(*Weak*)

The idea of the account is as follows. Since *Strong* and *Weak* are contextually equivalent, then they are both relevant (as the prejacent is always relevant and relevance is closed w.r.t. contextual entailment). Thus, the embedded implicature is mandatory. The sentence *every*(*Weak*) in (78b) has a global plain meaning that is logically stronger. Yet, because of the mandatory embedded implicature $\neg$*Strong*, this sentence (78b) is effectively equivalent to (79), with an overt *only* associated with *some* embedded in the restrictor of the universal quantifier. Oddness follows from the fact that, because of the embedded implicature or the embedded *only*, the restrictor of the universal quantifier is contextually empty, which makes the sentence either a contextual tautology or a presupposition failure (depending on whether it is indeed the case that universal quantifiers trigger a non-emptiness presupposition on their restrictor).[24]

(79) #Every father only SOME of whose children. . . , will pay a fine.

What about the sentence *every*(*Strong*) in (78a), that has a global plain meaning that is logically weaker? Why isn't it odd because of the matrix implicature corresponding to the matrix alternative *every*(*Weak*)? The reason is that this matrix alternative has a mandatory embedded implicature and is thus effectively equivalent to (79). Thus, this matrix alternative is a contextual contradiction or a presupposition failure. In particular, this matrix alternative is not in any way equivalent to the matrix prejacent *every*(*Strong*). The latter therefore triggers no mandatory matrix scalar implicature and thus no oddness is predicted. In the end, the reason why sentence (78b) sounds odd despite the fact that it has a strong global plain meaning is that it triggers a mismatching embedded implicature. The reason why sentence (78a) sounds fine despite the fact that it has a weak global plain meaning is that its potentially harmful matrix alternative is made inoffensive by its mismatching embedded implicature.

This account shows that the pattern of oddness observed in Section 4 falls into place, once oddness is computed also at the embedded level. Turning the perspective upside down, I conclude with the main point of this paper: the pattern of oddness in DE contexts uncovered in Section 4 provides new evidence for embedded scalar implicatures — actually, embedded where you would least expect them, namely in DE environments.

6.4 An additional argument

The two sentences (80) have the same plain meaning: how come that (80a) sounds fine while (80b) sounds odd in the context considered?[25] I would like to suggest that the generalization presented in Section 4 and the account just sketched shed some light on this puzzle.

(80) In this department, all professors get together at the end of the semester and decide a grade to assign to all of their students.

a. It is false that this year all professors assigned an A.

b. #This year, not all professors assigned an A.

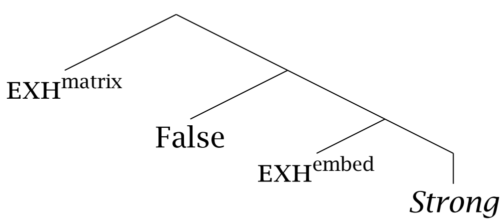
24 Thanks to Kai von Fintel for helping me clarify this point.

25 This example came up in conversation with Emmanuel Chemla.

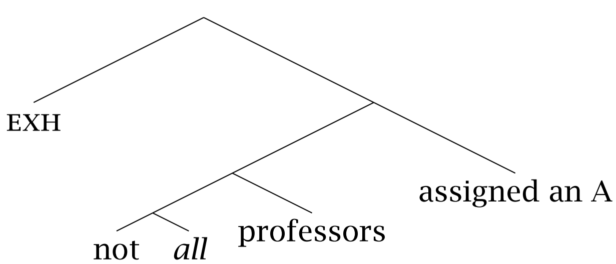
Sentence (80a) fits into the scheme $O_{\text{DE}}(\textit{Strong})$ considered in Section 4, where O_{DE} is the propositional operator *False* and the embedded constituent *Strong* is (81a). In the context considered here, all professors assigned the same grade, thus ensuring contextual equivalence between the embedded constituent *Strong* and the alternative *Weak* in (81b) obtained by replacing *all* with *some*. The left hand side of generalization (52) thus applies, and correctly predicts sentence (80a) to be fine.

(81) a. *Strong* = All professors assigned an A.
b. *Weak* = Some professors assigned an A.

Let me make the point a bit more explicit. In the case of sentence (80a), the strong scalar item *all* occurs in a clause embedded under the DE operator *False*. Thus, there is space for an embedded exhaustivity operator in between the DE operator and the embedded clause *Strong*, as in (82). Despite the fact that sentence (80a) has a globally logically weaker plain meaning than the alternative obtained replacing *Strong* with *Weak*, it is fine precisely because of this embedded exhaustivity operator (see the Appendix for details).

(82)


Sentence (80b) has the same plain meaning as sentence (80a), but has a very different syntactic structure. Plausibly, negation in (80b) is built inside the determiner phrase, as in (83). Thus, there is no space for an embedded exhaustivity operator in this case, and the sentence only has the matrix one. Since the prejacent of the matrix exhaustivity operator is logically weaker than the alternative with *all* replaced by *some* and since the matrix exhaustivity operator is the only one around, the sentence is correctly predicted to sound odd.

(83)


In conclusion, the contrast between the two sentences (80) is due to the fact that, despite having the same plain meaning, they have different syntactic structures, that imply a different distribution of exhaustivity operators. The contrast in (80) thus makes a case for theories of scalar implicatures that posit a syntactic trigger in the form of a covert but syntactically realized exhaustivity operator.

6.5 Nothing changes for plain cases

In Subsection 2.2, I stated my initial version of the obligatoriness of the exhaustivity operator, namely assumption (20) that the exhaustivity operator is mandatory at matrix level. Of course, there are plain cases that trigger no matrix implicature. Aren't those counterexamples to assumption (20)? In Subsection 2.4, I have argued they are not. When an implicature is absent, that is not because the exhaustivity operator is absent but rather because relevance has trimmed from its domain the corresponding alternative. Does this logic carry over to the extension from matrix scope (20) to arbitrary scope (75)? It is well known that scalar implicatures are usually unavailable or at least dispreferred in DE environments in plain cases (without focus on the scalar item). For instance, sentence (84a) can hardly be interpreted as saying that only students who did some *but not all* of the homework got an A, thus implying that those very diligent students who completed the assignment did not get an A. In other words, the existential quantifier in (84a) triggers no embedded implicature. Furthermore, the embedded implicature in (84a) remains unavailable even in cases where the corresponding alternative is mentioned in the previous discourse and thus plausibly made relevant, as illustrated in (84b).[26] How is the unavailability of implicatures in DE environments compatible with my assumption (75) that the exhaustivity operator is mandatory also in DE environments and that the oddness of (84c) is due to the fact that the existential quantifier *some children* triggers a mismatching implicature embedded in the DE environment provided by the restrictor of the universal quantifier *every father*?

(84) a. Every student who did some of the homework will get an A.
 b. A: I did all of the homework. Will I get an A?
 B: You definitely will. Everyone who did some of the homework will get an A.

26 Thanks to an anonymous S&P reviewer for pointing out this fact to me.

c. #Every father some of whose children have a funny last name will be fined.

In the rest of this Subsection, I argue that the logic used in Section 2.4 to make sense of mandatory matrix exhaustivity operators can indeed be extended to the case of mandatory embedded exhaustivity operators, thus making sense of these puzzles.

For concreteness, let me assume Fox & Spector's (2009) account for why scalar implicatures are dispreferred in DE environments in plain cases. They note that these scalar implicatures embedded in DE environments would make the global meaning weaker. They thus suggest that implicatures embedded in DE environments are ruled out by an *Economy Principle* that disfavors implicatures whenever they weaken the global meaning, rather than strengthening it. I have suggested that the exhaustivity operator is mandatory and that the availability of an implicature depends on whether the corresponding alternative belongs to the domain of the exhaustivity operator. Within this framework, the Economy Principle can thus be construed as condition (85) on licit domains of exhaustivity operators.

(85) An occurrence of the exhaustivity operator is ungrammatical whenever its domain contains an alternative whose corresponding implicature leads to a weakening of the global meaning.

The existential quantifier *some of the homework* in (84a) is associated with an exhaustivity operator embedded within the DE environment provided by the restrictor of the universal quantifier *every student*. The Economy Principle (85) demands as few relevant alternatives as possible in the domain of this embedded exhaustivity operator. In fact, any alternative that gets negated in that DE embedded environment will make the overall meaning weaker. The domain of the embedded exhaustivity operator has been defined in (16) as the intersection between the formal set of excludable alternatives and the set of relevant alternatives. In order to determine the domain of the embedded exhaustivity operator, we thus need to consult relevance. Suppose that the *all*-alternative has been mentioned in the previous discourse, as in (84b). Thus, its relevance has been increased. But the fact that previously mentioned alternatives are relevant is only a rule of thumb, not a rule of grammar. Some kind of local accommodation of relevance is thus possible in this case. For example, we might suppose that the discourse perspective has shifted in between the first one of B's sentences (*You definitely will*) in (84b) and the second one (*Everyone who did some…*). Only the first sentence really provides an answer to A's question, and thus shares with it the same perspective. With the second sentence, the discourse shifts to general rules of grading. From this new perspective, it does not matter whether a student has completed his homework. And the *all*-alternative can thus be construed as irrelevant at this point in the conversation, although it has indeed been held up as relevant up until now. As the domain of the exhaustivity operator is a subset of the set of relevant alternatives, this local accommodation of relevance gets the all-alternative out of the domain of the embedded exhaustivity operator and thus allows the Economy Principle (85) to be satisfied.

According to the proposal just sketched, the grammatical Economy Principle (85) imposes constraints on suitable contextual assignments to the relevance predicate $\mathcal{R}$. This is of course not surprising. For instance, gender features on free pronouns impose grammatical constraints on suitable contextual assignments to the variable introduced by the pronoun. Furthermore, the contrast in (86a) might provide some evidence that the Economy Principle (85) does indeed impose grammatical constraints on relevance.

(86) a. John did some of the homework.
b. Everyone who did some of the homework will get an A. John did some. Thus he will get an A.

Out of the blue, sentence (86a) triggers a robust not-all scalar implicature that John did not complete his assignment. This sets the baseline. Against this baseline, note that the same sentence embedded within the dialogue (86b) loses almost completely its not-all implicature. This is not surprising from the perspective I have just sketched. The first sentence *Everyone who did some...* in (86b) has an existential quantifier embedded in a DE environment. In order to satisfy the Economy Principle (85), the relevance predicate $\mathcal{R}$ that appears in this first sentence thus needs to be chosen to the effect that it is not relevant whether the entire homework has been completed. It makes sense to assume that, by default, the relevance predicate stays the same throughout the conversation. When we hit the second sentence *John did some* in (86b), it is therefore still irrelevant whether the homework has been completed. As the alternative *John did all* is not relevant, the corresponding implicature is thus correctly predicted to be unavailable.

Let me take stock. Implicatures embedded in DE environments are unavailable or at least dispreferred in plain cases. Following Fox & Spector (2009), I have assumed that this is due to an Economy Principle. Within the framework suggested in this paper, I have stated this principle as a condition on licit domains of exhaustivity operators, as in (85). As the domain of the exhaustivity operator is the intersection of the formally defined set of excludable alternatives and the contextually assigned set $\mathcal{R}$ of relevant alternatives, the Economy Principle (85) ends up imposing conditions on licit contextual assignments to relevance, just like gender features impose conditions on licit contextual assignments to free pronouns. In other words, it forces alternatives of scalar items embedded in DE environments to be analyzed (or perhaps reanalyzed through local accommodation) as not relevant. But crucially, such a local accommodation of relevance is not possible in the case of the odd sentence (84c). In this case, the embedded prejacent is (87a) and its alternative is (87b). Suppose that, in order fulfill the Economy Principle, we get the alternative (87b) out of the domain of the embedded exhaustivity operator by dooming it irrelevant. But how could that be?

(87) a. *Some* of x's children have a funny last name.
b. *All* of x's children have a funny last name.

The prejacent (87a) of the embedded exhaustivity operator is relevant, say because of axiom (19a) on relevance. Furthermore, the prejacent (87a) and the alternative (87b) say the same thing, express the same proposition, once we factor in the piece of common knowledge that all children inherit the last name of their father and thus all share the same last name. It is thus impossible for the alternative (87b) not to be relevant too, as axiom (19b) requires relevance to be closed w.r.t. contextual equivalence. In the case of (84a)-(84b), we could ensure that the embedded alternative was not relevant by isolating the sentence from the surrounding discourse through a local reanalysis of relevance, and thus satisfy the Economy Principle. But nothing like that helps in the case of the odd sentence (84c). In fact, what forces the alternative to be relevant in this case is the relevance of the prejacent, which lies well inside the sentence itself, not in the surrounding discourse. The only way to satisfy the Economy Principle in the case of (84c) would be to give up the piece of common knowledge that ensures the contextual equivalence between the prejacent (87a) and the alternative (87b). This is why the sentence mismatches with common knowledge.

7 Conclusion

In Magri 2009a, I argued that an implicature is mandatorily triggered when it corresponds to an alternative which happens to be contextually equivalent to the prejacent, leading to an effect of oddness. If that proposal is on the right track, then oddness can be used as a diagnostic to detect scalar implicatures. This is the perspective adopted in this paper. In particular, I have used this diagnostic in order to detect a new pattern of implicatures embedded in the scope of DE operators. In plain cases, these implicatures are not visible, say because an Economy Principle such as (85) rules them out, by dooming the corresponding alternative irrelevant. But in the case in which the embedded alternative is contextually equivalent to the embedded prejacent, the implicature can be detected through oddness. The paper has thus contributed new evidence to the recent debate on the existence of embedded scalar implicatures.

Appendix: The details of the account

In this Appendix, I illustrate in detail how the proposed account meets the two challenges outlined in Subsection 6.1.

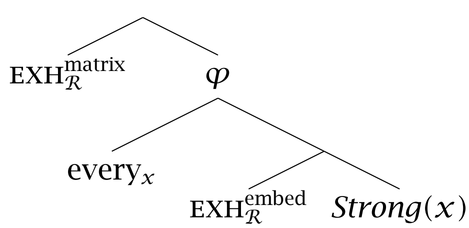
A.1 Meeting the first challenge raised by generalization (49)

Consider again the odd sentence (88a) of the IW pattern and the fine sentence (88b) of the IF pattern.

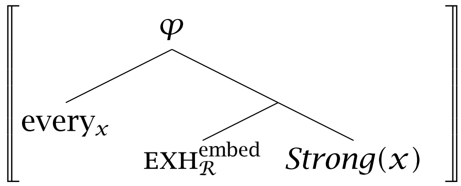
(88) a. #$\underbrace{\text{Italian women}}_{\textit{Strong}}$ come from a warm country.

b. $\surd$Every $\underbrace{\text{father whose children have a funny last name}}_{\textit{Strong}}$ must pay a fine.

By assumption (75) that the exhaustivity operator is mandatory at every scope site, the LF of sentences (88) is (89), with both a matrix and an embedded exhaustivity operator. Here, *Strong* is as in (39) and (45), respectively.

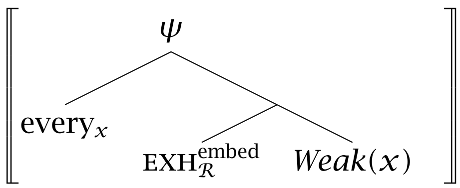
(89)


Since *Weak* is not excludable w.r.t. *Strong*, the embedded exhaustivity operator does nothing and the matrix prejacent φ ends up with the truth conditions (90).

(90) 

$= \text{every}_x\ Strong(x)$

The matrix alternative is ψ obtained by replacing *Strong* with *Weak*, where *Weak* is as in (39) and (45), respectively. Depending on whether *Strong* is relevant or not, we get one of the two truth conditions in (91). In either case, the alternative ψ is logically excludable w.r.t. the prejacent φ.

(91) 

$=$

$$= \begin{cases} \text{every}_x Weak(x) & \text{if } Strong \notin \mathcal{R} \quad \text{case (a)} \\ \text{every}_x (Weak(x) \text{ and not } Strong(x)) & \text{if } Strong \in \mathcal{R} \quad \text{case (b)} \end{cases}$$

Consider first the case of (88a), characterized by the fact that *Strong* and *Weak* are *not* contextually equivalent. No matter whether the alternative ψ ends up with the truth conditions (91a) or (91b), it is contextually equivalent to the prejacent φ in (90). Axioms (19) on relevance thus force the matrix implicature in place. In conclusion, we derive the truth conditions $[\![(88\text{a})]\!] = \varphi \wedge \neg\psi$, that are a contextual contradiction.[27] Consider next the

27 I am assuming a framework where there are exhaustivity operators both in the prejacent as well as in the alternatives, as in (90) and (91), raising the issue of whether the Economy Principle (85) applies only to the exhaustivity operators in the prejacent or also in the alternatives. If the Economy Principle (85) applies also to the alternative (91), then it rules out the option of truth conditions (91b) for the alternative ψ, leaving open only the option of truth conditions (91a). If instead the Economy Principle (85) does not apply also to the alternative (91), then both truth conditions (91a) and (91b) are in principle available for the alternative ψ. In the latter case, it is crucial to choose the definition of excludable alternatives in (15) in terms of non-contradictoriness, rather than the more traditional one that requires a scalar alternative to asymmetrically entail the prejacent in order to qualify as excludable. In fact, if the truth conditions of the alternative ψ are (91b), then ψ does not logically asymmetrically entail the prejacent φ in (90). Thus, if I had assumed the alternative definition of the set of excludable alternatives in terms of asymmetric entailment, the matrix exhaustivity operator would have ended up with no alternatives. With the result that no

case of (88b), characterized by the fact that *Strong* and *Weak* are *indeed* contextually equivalent. By axioms (19) on relevance, *Strong* is relevant, because contextually equivalent to the embedded prejacent *Weak*. The embedded implicature is forced in place and the truth conditions of the alternative ψ are therefore (91b). Because of the contextual equivalence between *Strong* and *Weak*, the alternative $\psi = every_x(Weak(x)\ and\ not\ Strong(x))$ is thus either a contextual tautology (because the universal quantifier is restricted by a contextually empty restrictor) or a presupposition failure (if universal quantifiers trigger a non-emptyness presupposition for their restrictors). In either case, the matrix alternative ψ is not contextually equivalent to the matrix prejacent φ in (90). Thus, no matrix implicature is predicted, and we end up with the fine truth conditions $[\![(88\text{b})]\!] = \varphi$.[28]

A.2 Meeting the second challenge raised by generalization (49)

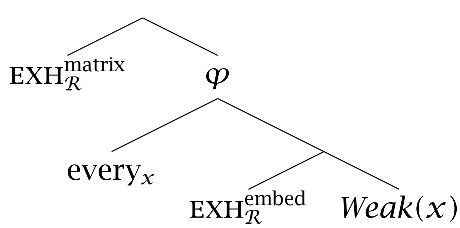
Consider again the good sentence (92a) of the IW pattern and the odd sentence (92b) of the IF pattern.

(92) a. $\surd\underbrace{\text{Italians}}_{Weak}$ come from a warm country.

b. #Every $\underbrace{\text{father some of whose kids have a funny last name}}_{Weak}$ will be fined.

By assumption (75) that the exhaustivity operator is mandatory at every scope site, the LF of sentences (92) is (93), with both a matrix and an embedded exhaustivity operator. Here, *Weak* is as in (39) and (45), respectively.

(93)



mismatching implicature would have been derived. This difficulty does not arise with the definition (15) of the set of excludable alternatives in terms of non-contradictoriness: in this case, no matter whether the alternative ψ ends up with the truth conditions (91a) or (91b), its negation is logically compatible with the prejacent φ in (90) and thus counts as a matrix excludable alternative.

28 Another option is of course that the alternative $\psi = every_x(Weak(x)\ and\ not\ Strong(x))$ in (91b) is ruled out by the Economy Principle (85), provided that the latter applies also to the exhaustivity operators in the alternatives.

Depending on whether *Strong* is relevant or not, the prejacent φ has one of the two truth conditions in (94).

(94) $$\left[\!\!\left[[_{\varphi}\ \text{every}_x\ [\ \text{EXH}^{\text{embed}}_{\mathcal{R}}\ \mathit{Weak}(x)\ ]\ ] \right]\!\!\right] =$$

$$= \begin{cases} \text{every}_x \mathit{Weak}(x) & \text{if } \mathit{Strong} \notin \mathcal{R} \quad \text{case (a)} \\ \text{every}_x (\mathit{Weak}(x) \text{ and not } \mathit{Strong}(x)) & \text{if } \mathit{Strong} \in \mathcal{R} \quad \text{case (b)} \end{cases}$$

The matrix alternative is ψ obtained by replacing *Weak* by *Strong*, where *Strong* is as in (39) and (45), respectively. Since *Weak* is not excludable w.r.t. *Strong*, the embedded exhaustivity operator in the alternative ψ does nothing. Thus, ψ ends up with the plain truth conditions (95).

(95) $$\left[\!\!\left[[_{\psi}\ \text{every}_x\ [\ \text{EXH}^{\text{embed}}_{\mathcal{R}}\ \mathit{Strong}(x)\ ]\ ] \right]\!\!\right] = \text{every}_x\ \mathit{Strong}(x)$$

Consider first the case of (92a), where the two restrictors *Strong* and *Weak* are *not* contextually equivalent. Thus, I am free to assume *Strong* is not relevant, without violating the axioms (19) on relevance. In this case, the embedded exhaustivity operator in the matrix prejacent φ does nothing and the matrix prejacent φ ends up with the plain truth conditions (94a). Of course, the alternative ψ in (95) is not excludable w.r.t. the prejacent φ, and thus also the matrix exhaustivity operator does nothing. We thus derive the plain truth conditions $[\![(92\text{a})]\!] = \mathit{every}_x \mathit{Weak}(x)$, that are of course not a contextual contradiction. Consider next the case of (92b), where the two restrictors *Strong* and *Weak* are *indeed* contextually equivalent. By axioms (19) on relevance, *Strong* is relevant, because contextually equivalent to the embedded prejacent *Weak*. Thus, the truth conditions of the matrix prejacent are (94b). No matter what the matrix exhaustivity operator does, the matrix prejacent is thus ruled out, either because it is a presupposition failure or a contextual tautology (as the restrictor of *every* is contextually empty) or because it violates the Economy Principle (85).

References

Blutner, Reinhard. 1998. Lexical pragmatics. *Journal of Semantics* 15(2). 115–162. http://dx.doi.org/10.1093/jos/15.2.115.

Blutner, Reinhard. 2000. Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17(3). 189–216. http://dx.doi.org/10.1093/jos/17.3.189.

Chemla, Emmanuel & Benjamin Spector. 2011. Experimental evidence for embedded implicatures. *Journal of Semantics* 28(3). 359–400. http://dx.doi.org/10.1093/jos/ffq023.

Chierchia, Gennaro. 1995. Individual-level predicates as inherent generics. In Gregory N. Carlson & Francis Jeffry Pelletier (eds.), *The generic book*, 176–223. Chicago, IL: The University of Chicago Press.

Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures, Volume 3*, 39–103. Oxford: Oxford University Press.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. to appear. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Paul Portner, Claudia Maienborn & Klaus von Heusinger (eds.), *Handbook of semantics*. Berlin, New York: Mouton de Gruyter. http://semanticsarchive.net/Archive/WMzY2ZmY/CFS_EmbeddedSIs.pdf.

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and implicature in compositional semantics*, 71–120. New York: Palgrave Macmillan.

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29(5). 537–586. http://dx.doi.org/10.1007/s10988-006-9004-4.

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

Fox, Danny & Benjamin Spector. 2009. Economy and embedded exhaustification. Handout from a talk at Cornell, available at http://web.mit.edu/linguistics/people/faculty/fox/Fox_Spector_Cornell.pdf.

Geurts, Bart & Nausicaa Pouscoulous. 2009. Embedded implicatures?!? *Semantics and Pragmatics* 2(4). 1–34. http://dx.doi.org/10.3765/sp.2.4.

Grice, Paul. 1975. Logic and conversation. In Peter Cole & Jerry Morgan (eds.), *Syntax and semantics 3: Speech acts*, 41–58. New York: Academic: Academic Press.

Groenendijk, Jeroen & Martin Stokhof. 1984. *Studies on the semantics of questions and the pragmatics of answers*. University of Amsterdam PhD thesis. http://dare.uva.nl/en/record/123669.

Hawkins, John A. 1991. On (in)definite articles: implicatures and (un)grammaticality prediction. *Journal of Linguistics* 27(2). 405–442. http://dx.doi.org/10.1017/S0022226700012731.

Heim, Irene. 1991. Artikel und Definitheit. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, 487–535. De Gruyter: Berlin.

Horn, Laurence R. 1972. *On the semantic properties of logical operators in English*. UCLA PhD thesis.

Horn, Laurence R. 2005. The border wars: a neo-Gricean perspective. In Ken Turner & Klaus von Heusinger (eds.), *Where semantics meets pragmatics* (Current Research in the Semantics/Pragmatics Interface 16), 21–48. Elsevier.

Hurford, James R. 1974. Exclusive or inclusive disjunction. *Foundations of Language* 11(3). 409–411. http://www.jstor.org/stable/25000785.

Ippolito, Michela. 2008. On the meaning of *only*. *Journal of Semantics* 25(1). 45–91. http://dx.doi.org/10.1093/jos/ffm010.

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

Kratzer, Angelika. 1995. Stage-level and individual-level predicates. In Gregory N. Carlson & Francis Jeffry Pelletier (eds.), *The generic book*, 125–175. Chicago, IL: The University of Chicago Press.

van Kuppevelt, Jan. 1996. Inferring from topics. *Linguistics and Philosophy* 19(4). 393–443. http://dx.doi.org/10.1007/BF00630897.

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

Magri, Giorgio. 2009b. *A theory of individual-level predicates based on blind mandatory scalar implicatures: Constraint promotion for Optimality Theory*. MIT PhD thesis. http://dx.doi.org/1721.1/55182.

Magri, Giorgio. 2009c. A theory of individual level predicates based on blind mandatory scalar implicatures (extended version). MIT manuscript; available at http://web.mac.com/giorgiomagri/Site/home.html.

Magri, Giorgio. 2009d. Mismatching scalar implicatures. In Paul Egré & Giorgio Magri (eds.), *Proceedings of the MIT-Paris workshop on implicatures and presuppositions* (MITWPL 60), 153–168. Cambridge, MA: MIT Working Papers in Linguistics.

Magri, Giorgio. 2011. Another argument for embedded scalar implicatures based on oddness in downward entailing contexts. *Semantics and Linguistic Theory (SALT)* 20. 564–581. http://elanguage.net/journals/index.php/salt/article/view/20.564.

Maienborn, Claudia. 2004. A pragmatic explanation of the stage level/individual level contrast in combination with locatives. *Western Conference of Linguistics (WECOL)* 15. 158–170. http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:hebis:30-1111368.

Musan, Renate. 1997. Tense, predicates, and lifetime effects. *Natural Language Semantics* 5(3). 271–301. http://dx.doi.org/10.1023/A:1008281017969.

Percus, Orin. 1997. *Aspects of* A. MIT PhD thesis. http://dx.doi.org/1721.1/10235.

Percus, Orin. 2006. Antipresuppositions. In Ayumi Ueyama (ed.), *Theoretical and empirical studies of reference and anaphora: toward the establishment of generative grammar as an empirical science*. Report of the Grant-in-Aid for Scientific Research (B), Project No. 15320052, Japan Society for the Promotion of Science, 52-73.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17(3). 501–557. http://www.jstor.org/stable/4178501.

van Rooij, Robert. 2001. Conversational implicatures. *Proceedings of the 2nd SIGdial Workshop on Discourse and Dialogue*. 1–10. http://dx.doi.org/10.3115/1118078.1118102.

Sauerland, Uli. 2002. The present tense is vacuous. *Snippets* 6. http://www.ledonline.it/snippets/allegati/snippets6004.pdf.

Sauerland, Uli. 2003. A new semantics for number. *Semantics and Linguistic Theory (SALT)* 13. 258–275. http://www.elanguage.net/journals/index.php/salt/article/view/13.258.

Sauerland, Uli. 2004a. A comprehensive semantics for agreement. Paper presented at the "Phi-Workshop", McGill University, Montreal, Canada. http://homepage.mac.com/yasa/uli/agreement.pdf.

Sauerland, Uli. 2004b. On embedded implicatures. *Journal of Cognitive Science* 5(1). 107–137.

Sauerland, Uli. 2004c. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27. 367–391. http://dx.doi.org/10.1023/B:LING.0000023378.71748.db.

Sauerland, Uli. 2008. Implicated presuppositions. In Anita Steube (ed.), *The discourse potential of underspecified structures*, 581–600. Berlin, Germany: Mouton de Gruyter.

Schlenker, Philippe. 2011. *Maximize Presupposition* and Gricean reasoning. Manuscript, IJN and NYU. https://files.nyu.edu/pds4/public/MaxPres.pdf.

Schwarzschild, Roger. 1996. *Pluralities*. Dordrecht, Boston, London: Kluwer Academic Publisher.

Singh, Raj. 2009. *Maximize Presupposition!* and informationally encapsulated implicatures. *Sinn und Bedeutung* 13. http://elib.uni-stuttgart.de/opus/volltexte/2009/4360/pdf/sub13_band2_190609.pdf#page=233.

Spector, Benjamin. 2006. *Aspects de la pragmatique des operateurs logiques*. Universite Paris 7 PhD thesis. http://lumiere.ens.fr/~bspector/THESE_SPECTOR/THESE_SPECTOR_AVEC_ANNEXE2.pdf.

Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: on higher-order implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presuppositions and implicatures in compositional semantics*, 243–281. Palgrave-Macmillan.

Stalnaker, Robert C. 2002. Common ground. *Linguistics and Philosophy* 25. 701–721. http://dx.doi.org/10.1023/A:1020867916902.

Zondervan, Albert Jan. 2010. *Scalar implicatures or focus: An experimental approach*. University of Utrecht PhD thesis. http://igitur-archive.library.uu.nl/dissertations/2010-0517-200229/UUindex.html.

Giorgio Magri
IJN, ENS, CNRS
magrigrg@gmail.com