Numerals and scalar implicatures

Daniele Panizza, Gennaro Chierchia

Abstract

In this paper we explore the relation between the meaning of number denoting determiners ('numerals') and the polarity of the context in which they occur. We claim that when numerals are embedded in positive (i.e. Upward Entailing) contexts are given an upper bounded ('exactly') reading more often than when they are embedded in minimally different Downward Entailing contexts. For this mirrors the behavior of scalar triggers, we suggest that the stronger interpretation of a numeral is really due to a Scalar Implicature. We review the outcome of two experiments, a questionnaire and a reading task where eye movements were recorded, which tested empirically this claim. Finally, we discuss these findings in light of the current approaches on the semantics of numerals, analyzing in detail the relation between numerals meaning and scalar strenghtening.

1. Introduction

Sentences quantified by numerals may have (at least) two different meanings: an 'exactly' (or upper bounded) meaning and an 'at least' (or lower bounded) one. They are exemplified by (1a) and (1b) respectively:

(1) a. If I get in the Summer competition, I'll buy four golf clubs
b. If I buy four golf clubs, I'll never use my loose old clubs again

We feel that (1a)'s most salient reading is an upper bounded one (…, I'll buy four golf clubs and no more), while (1b) is more naturally interpreted as lower bounded (If I buy four golf clubs or more,…). This paper is devoted to arguing that this 'feeling' is both general (across speakers) and grounded in grammar and processing.

More specifically, although the interpretation of numerals is influenced by extra-linguistic factors (e.g. the utterance context, speaker and hearer's intentions, etc.) in the present paper we argue that there are also purely structural factors (i.e. the entailment pattern of the construction in which the numeral is embedded, that is, the polarity of the context) that systematically affect the distribution of lower bounded vs. upper bounded readings of numerals. In (1a) the numeral 'four' occurs in the consequent of a conditional. This position licenses entailments from subsets to supersets, and is therefore an 'Upward Entailing' (UE) environment. I.e. (2a) logically entails (2b) (but not vice versa):

(2) a. If I go to the pub, I'll drink a stout beer.
b. If I go to the pub, I'll drink a beer.

In (2), specifically, the entailing inference goes from the 'set of *stout beer* drinkers' to the 'set of *beer* drinkers', with the latter a superset of the former. The antecedents of conditionals, on the other hand, licence the opposite entailment pattern and hence constitute 'Downward Entailing' (DE) contexts. Sentence (3a) logically entails sentence (3b) (but not viceversa):

(3) a. If I drink a beer, I'll get a headache.
b. If I drink a stout beer, I'll get a headache.

The UE/DE contrast has been shown to be relevant to many grammatical phenomena (e.g. the distribution of negative polarity items), and is also relevant, we claim, to the distribution of the 'exact' vs. 'at least' construals of numerals. Here is, specifically, the theoretical claim we put forward in the present paper:

(4) Claim:
The upper bounded interpretation of numerals occurs preferentially in UE contexts with respect to minimally different DE contexts.
The lower bounded interpretation of numerals occurs preferentially in DE contexts with respect to minimally different UE contexts.

Two UE and DE contexts are minimally different in case they share the same lexical material with the exception of the replacement of a DE functor for an UE one, or vice versa.

This claim does not say anything about the absolute proportion of upper/lower bounded interpretations intended by the speaker in any context. It is well known, for instance, that contextual (i.e. knowledge or discourse context based) factors may affect the interpretation of a scalar item (cf. Breheny, Katsos, & Williams, 2006). Our thesis posits that, whatever the base rate of interpreting a numeral with the upper or lower bounded reading is, structural grammatical factors, like the entailing property of the proposition, systematically affect readers' interpretation as stated in (4). Clearly, native speakers' intuitions of the type often used in current linguistic theorizing are insufficient to establish (4), as the frequency of the upper/lower bounded interpretation of numerals may vary across speakers, or even within a single one. For this reason we will discuss, in the present paper, an experimental study (cf. Panizza, Chierchia and Clifton, in press) that we run to specifically test the claim pointed out in (4). As for processing concerns, (4) does not say anything particular about the cost of deriving an upper/lower bounded reading in either a UE or a DE context. However, many studies in psycholinguistics attest that a dispreferred interpretation of a word or a sentence imposes a processing load, with respect to the preferred one (cf. Rayner, 1998, for a review of research using the eyetracking methodology we will use). We return to the question of processing cost after a discussion of the interpretive preferences for numerals.

The structure of the paper is as follows. We first discuss, in section 2 and 3, some competing approaches on the meaning of numerals. Then, in sections 4 and 5, we present two experiments investigating the offline and online behaviour, respectively, of subjects while they are presented with sentences containing numerals. Finally, in section 6 we discuss how our findings affect the theoretical debate reviewed in sections 2 and 3.

2. An overview of theoretical approaches to Scalar Implicatures

Claims similar to (4) have been put forth in the literature, especially in the context of work on Scalar Implicatures (SIs; Grice, 1989). Consider the variation in interpretation of items like *some*, displayed in (5a,b):

(5) a. The professor saw some of his students and he'll go out for dinner with them.
b. If the professor saw some of his students, he'll go out for dinner with them.
c. The quantifier scale: <some, many, most, every>

Sentences (5a) and (5b) differ minimally in form and meaning. Yet, in sentence (5a) *some* seems to mean 'some though not all' (upper bounded interpretation), while clearly that isn't how *some* is interpreted in (5b) (lower bounded interpretation). According to Grice and the Neogricean literature (e.g. Levinson, 2000; Horn, 1972, 2007), this phenomenon is discussed in terms of SIs. Sentences involving *some* are often considered against the background of alternatives constituted by other quantifiers that might be relevant, such as those in (5c). Use of *some* prompts the hearer to assume that the stronger alternatives do not hold (on the basis of simple assumptions on how conversation proceeds – the Gricean maxims). The same phenomenon is argued to affect the interpretations of other scalar items like connectives (*or*, *and*), modals (*may, must*), gradable adjectives (like *warm*, *hot*), etc. Within the Neogricean tradition, the 'some and possibly all' meaning of *some* is taken to be basic, and the denial that a stronger interpretation of *some* holds (which results in the 'some but not all' reading) is analyzed as a scalar implicature (SI).

An alternative development of Grice's insights on these matters, namely Relevance Theory (Sperber & Wilson 1986, 1995), maintains that interpretive alternations such as those in (5a,b) arise in a different manner. The basic meaning of *some* is lower bounded ('at least some'), and inferences that specify such a meaning towards the upper bounded construals come about only if it is required to maximize relevance. 'Relevance' is a property of the situation and the stimulus (a sentence, in this case). To arrive at the intended meaning, the hearer puts his or her whole cognitive domain (perceptions, world knowledge, intentions, emotions etc.) in relation with the sentence. If the hearer's expected relevance is satisfied with *some* being interpreted with its basic (i.e. lower bounded) meaning, additional effort will not be invested in drawing an inference towards the upper bounded meaning. This approach seems designed to predict the interpretations of cases like (6) vs. (7).

(6) Mary, who is a big eater, will surely eat some of that cake.
(7) Mary, who is scared of dentists, will surely have to spend some of her time tomorrow at the dentist.

Some in (6) might have greater chances to be interpreted in a lower bounded way than in (7) given the context and the nature of the task. We will discuss the implications of this approach with respect to the meaning of numerals in the next pages.

More recently, Chierchia (2004) and Chierchia, Fox and Spector (in press) have argued for the view that scalar implicatures are really part of core grammar. According to them, scalar implicatures arise through mechanisms analogous to those that drive association with focus (cf. Rooth, 1985). Building more closely on the insights of the Neogriceans, Chierchia et al. (in press) propose to represent strengthened (upper-bounded, scalar implicature) meanings grammatically by adding a silent operator O, with an effect similar to that of focus sensitive operators like *only*. Consider (8).

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

Sentence (8a) may or may not receive an upper bounded interpretation; sentence (8b) must be interpreted in an upper bounded way. The word *only* can be said to overtly 'exhaustify' the proposition, making it incompatible with any stronger interpretation of the sentence (see Chierchia et al., in press, for extended discussion).

Exhaustification can be covert as well as overt. Consider (9):

(9) a. i. Speaker A: So did you see the students?
ii. Speaker B: Well, I saw MARY
b. O [I saw Mary] where O = only

Speaker B's utterance, in which *Mary* is focused, is clearly interpreted as equivalent to 'I saw only Mary'. This might be understood by assuming that there is a covert counterpart of only that is used in interpreting (9a.ii), as schematically indicated in the Logical Form (9b). Clearly, the kind of covert exhaustification that must be at play in cases like (9) might well be responsible for the upper bounded interpretation of sentences like (8a). According to this view, the lower bounded interpretation would correspond to the (unexhaustified) Logical Form in (10a); the upper bounded one would correspond to (10b).

(10) a. some homework$_x$[John graded x]
b. O[some homework$_x$[John graded x]]

The Logical Form in (10b) would be interpreted just like (8b), giving rise to the upper bounded construal.

One prediction that sets aside the Grammatical View from Neogriceans or relevant theorists concerns the possibility of embedded implicatures. On the Grammatical View, where implicatures arise through a covert counterpart of *only*, one expects there to be embedded implicatures, for such an operator will be able to occur in embedded positions (just like its overt counterpart). On the other hand, in approaches where implicatures arise through a 'global', post-compositional process, involving speakers' intentions and the like, it is hard to see how implicatures could ever enrich the meaning of embedded constituents (thereby affecting the compositional part of semantics). This problem (acknowledged in, e.g., Levinson, 2000) along with the different predictions of the two families of approaches is discussed at length in Chierchia et al. (in press).

3. Numerals and Scalar Implicatures

Much of the analysis of SIs has involved quantifiers such as *some*. Whether the analysis applies to numerals, which are clearly scalar terms, is controversial. For example, Horn (1992, 2004) argues that quantifiers and numerals behave differently under negation. Moreover, it has also been argued that numerals can also get an 'at most' reading, which is not available with other scalar quantifiers and could not be derived as a SI. Carston (1998) and Musolino (2004) present examples of the following sort in favor of the existence of an 'at most' reading.

(11) If you miss two shots, you will still win

We believe this claim to be factually wrong. The appearance of an 'at most' construal for the numeral in (11) has, we think, a different account. Sentences like (11) have an 'exactly' interpretation (something we independently know to be possible), *plus* a presupposition (in the case at hand, triggered by 'still') that can be paraphrased as

(12) If you miss no or exactly one shot, you will win.

Sentences (11) and (12) together are equivalent to 'at most'. Be that as it may, the main point we want to make in this paper is orthogonal to the (non) existence of 'at most' readings for numerals, and concerns the distribution of the 'at least' and the 'exactly' reading.

Huang and Snedeker (2009) conducted an experimental study that speaks to this issue. In this study eye movements of subjects were recorded while they attended to a visual scenario depicted on a monitor and listened to some sentences (visual world paradigm). These sentences contained numerals like *two* and *three* and scalar quantifiers like *some* and *all*. Subjects had to pick the character, present in the scenario, who was described by the sentence they heard. The authors found a delay in looking toward the referent of a phrase containing *some* (e.g. "point to the boy with some of the socks", where the scenario dysplayed a boy with two socks, a boy with three soccer balls, a girl with two socks and another girl with no objects) compared to unambiguous phrases containing *all* and phrases containing the numerals *two* or *three* (where *three* was unambiguous but *two* could arguably have a 'two or more' meaning). They argued that resolving *some* as 'some but not all' incurred into higher processing costs compared to unambiguous sentences and sentences with possibly ambiguous numerals.

One goal of our study was to determine whether the generalization in (4) holds for numerals as much as it seems to hold for quantifiers and other scalar items. For our purposes, we may briefly classify the positions on the semantics of numerals in four main families, schematically laid out in (13):

(13) a. Lexicalist approaches. Numerals are lexically ambiguous between two construals.
(Horn, 1992)
b. Numbers are underspecified. Pragmatic enrichment is driven by relevance.
(Relevance Theory: Sperber and Wilson, 1986, 1995; Carston, 1998)
c. Numbers are exact. The 'at least' reading comes through by a pragmatic operation of existential closure.
(Breheny, 2008)
d. Scalar Implicatures. The upper bounded construal of numerals is a SI derived on the basis of their lower bounded construal.
Approaches of this third type divide further into two streams:
i. Purely pragmatically based (Neogricean; cf. Kadmon, 2001)
ii. Grammatically based (Chierchia, 2004; Chierchia, Fox and Spector, in press)

Let us now comment briefly on each proposal in turn. Concerning position (13a), typical lexical ambiguities are not sensitive to the entailing properties of the context in which a lexical item is inserted. For example, the interpretation of the words like 'bank' or 'run' will typically remain constant across (14a) or (14b) even though they occur in a UE or a DE context, respectively.

(14) a. i. If it isn't too crowded, you'll like that bank
ii. If you like that bank, you'll go there often
b. i. If you pray, the car will run
ii. If the car will run, we are in luck
[run = functioning/ vs. go fast, or partake in a competition]

Notice that this holds even if the relation among different meanings of a word do display subsets/superset relations, and hence entailment would be potentially relevant (if a car runs in a competition, it has to be running - in the sense of functioning - but not viceversa). So if (4) is right, lexicalist approaches would be faced with the task of explaining why the lexical ambiguity of numbers turns out to be sensitive to entailment patterns while other lexically ambiguous words are not (even when they potentially could).

In so far as underspecification analyses in (13b) are concerned, the point of contention is the recognition of a purely structural factor in the emergence of the lower bounded vs. upper bounded contrast in numeral interpretation. If (4) above is right, then the entailment characteristic of two minimally different local environments of the numeral would affect our interpretation, regardless of any relevant contextual factor like world knowledge, discourse context, speaker's intentions, etc. We do not see how the UE/DE context could be claimed to affect relevance (apart from building into relevance a condition equivalent to (4)).

Breheny (2008) may be viewed as a representative of theories of type (13c). It might be useful to sketch here a simplified approach loosely inspired to his ideas (as Breheny's own proposal cannot be summarized within the bounds of the present work). One might take (15b) as the basic interpretation of (15a):

(15) a. John loves two cats
b. $| cat_D \cap \{x: John\ loves\ x\} | = 2$

Let us explain (15b) in more detail. Formula (15b) says that the intersection of the set of cats in a domain D with the set of things that John loves has cardinality two. Such a formula is true if and only if John loves exactly two cats in D. It therefore appears to be an accurate rendering of the upper bounded construal of sentence (15a). D in (15b) is a variable over the domain of discourse. This is meant to represent the fact that we may utter (15b) having a specific domain in mind (e.g. the cats that live in John's neighbourhood). Such a domain is usually contextually specified. However, conceivably, some contexts, for a variety of reasons, may drive a process of existential closure of such domain variable, which would result in something like:

(16) $\exists D| cat_D \cap \{x: John\ loves\ x\} | = 2$

This formula says that there is a way of picking a domain D which would make (15a) true. For this to be so, it has to be the case that for some domain D, John loves exactly two cats in D (which is of course compatible with John loving more than two cats). Formula (16) has, in other words, the truth conditions characteristic of the lower bounded reading. This illustrates one way of adopting as basic the 'exact' reading of numerals and deriving through a semantic/pragmatic process the 'at least' one.

Our claim (4) could be compatible with theories of the type (13c), *if* one is willing to add to them the condition that an operation like (16) is performed preferentially in DE contexts than UE ones. This may be difficult to reconcile with the view that (16) is a purely *pragmatic* inference, for it is not clear why (or how) a pragmatic tendency should be sensitive to whether numerals are embedded in UE vs. DE contexts. If, on the other hand, the existential closure of domain is an option provided by grammar (much like different scope options), it would be quite natural to maintain that such an operation might be subject to a processing constraint that links its

preferential occurrence to contexts in which this operation leads to stronger (i.e. more informative) statements.[1]

Finally, recall that the main assumption of SI-approaches such as (13d) is that the lower bounded interpretation constitutes in some sense the core meaning of sentences with numerals, with the upper bounded one derived as an implicature.[2] Generalization in (4) is a priori compatible with such an assumption. However, it also carries some implications that seem to favour the grammatical approach (13d.ii) over the purely pragmatic one (13d.i). Consider the natural explicit paraphrases of the upper vs. lower bounded readings of sentences in (17) and (18), with the upper bounded ('exactly') meanings obtained by adding an operator (O) either in the consequent (17b) or in the antecedent (18b) of a conditional clause.

(17) a. If I get in the summer competition, I'll buy (at least) four golf clubs
b. If I get in the summer competition, **O**[I'll buy four golf clubs]
= If I get in the summer competition, I'll buy exactly four golf clubs

(18) a. If I buy (at least) four golf clubs, I'll never use my loose old clubs again
b. If **O**[I buy four golf clubs], I'll never use my loose old clubs again
= If I buy exactly four golf clubs, I'll never use my loose old clubs again

On a SI approach, (17a) and (18a) would correspond to the core ('at least') meaning, while (17b) and (18b) would arise as implicatures (when warranted). In other words (17a) would be the core meaning of (1a) and (17b) would be the implicature enriched reading of (1a), and similarly for (18a,b) with respect to (1b). A possible account of claim (4), consistent with the SI approach, is to appeal to a principle of the following sort:

(19) Optimize Informativeness: Preferably, embed an implicature in contexts where it leads to a stronger statement than its alternative without the implicature.

The nature of Optimize Informativeness should be fairly clear: sentence (17b) (i.e. the upper bounded reading) entails sentence (17a) (i.e. the lower bounded reading). Namely, (17b) is logically stronger (and hence more informative) than (17a). On the other hand, sentence (18a) (the lower bounded reading) logically entails (18b) (the upper bounded reading). Namely (18a) is logically stronger (and hence more informative) than (18b). In other words, embedding an implicature in an UE context leads to strengthening (with respect to the statement without the implicature), while embedding an implicature in a DE context leads to weakening. This is a general property of DE environments: they reverse the entailment patterns of UE environments. So the point is that if we embed an implicature, we prefer to do it when this leads to strengthening with respect to the sentence without the implicature, as exemplified in (20) (where the arrow stands for the entailment, i.e. strength, relation).

[1] However, on such an approach one would have to explore whether existential closure of domains is possible with other quantifiers, and if not, why not.

[2] There is an important caveat to be made in this connection. Saying that the basic meaning of a *sentence* containing a numeral is the lower bounded interpretation doesn't entail that the *lexical* meaning of the numeral itself is the lower bounded one. It is perfectly conceivable that the lexical meaning of a numeral is the exact/upper bounded one and that the lower bounded interpretation arises as part of the compositional semantics of sentences (see e.g. Landman, 2003, for an approach along these lines).

(20) UE: 'at least four' ← 'exactly four'
DE: 'at least four' → 'exactly four'

While (19) is a natural, if sophisticated, principle, it is not obviously compatible with the standard Neogricean views, if for no other reason, that such an approach cannot motivate smoothly the very existence of embedded implicatures.

In conclusions, both approaches of the type in (13c) and (13d) are consistent with our proposed generalization (4), if the process that drives the derived interpretation of numerals is subject to a principle sensitive to logical strength such as (19).

It might be worth underscoring that the generalization in (4) says that (17b) as interpretation of (1a) is preferred to (18b) as interpretation of (1b). This specifically means that (17b) should occur more often as an interpretation of (1a) than (18b) as an interpretation of (1b). What (4) says *nothing* about is whether the addition of an implicature to a core meaning is per se a costly process. We address this issue in discussing Experiment 2, which investigated the on-line processing costs of sentences with numerals.

Summing up, our goal is to investigate whether claim (4) is true or not. If it is, this will have rich consequences for our understanding of how numerals are interpreted, and a host of related issues at the interface between grammar, pragmatics and processing. It would be moreover quite striking to find out that we spontaneously submit to a rather abstract logical regularity like the one (4) relies on.

4. Experiment 1: an off line semantic judgment test

In this experiment, we explicitly asked 48 undergraduate students to indicate their interpretation of the numeral determiner. They had to choose between the stronger *exactly* meaning and the weaker *at least* one by checking the appropriate box, after reading a sentence of the kinds displayed in (21) and (22).

(21a) Giovanni ha **due** macchine in garage e parcheggia una motocicletta nel cortile esterno.
John has **two** cars in the garage and he parks a motorcycle in the courtyard.
(21b) Se Giovanni ha **due** macchine in garage, parcheggia una motocicletta nel cortile esterno.
If John has **two** cars in the garage he parks a motorcycle in the courtyard.
(22a) Nel mio quartiere ogni ragazza ha **due** fratelli più grandi e desidera una sorellina di età inferiore.
In my neighborhood every girl has **two** older brothers and she wishes a younger sister.
(22b) Nel mio quartiere ogni ragazza che ha **due** fratelli più grandi desidera una sorellina di età inferiore.
In my neighborhood every girl who has **two** older brothers wishes a younger sister.

As can be observed, the sentences in (21) and (22) differ minimally. Items (21a) and (21b) correspond to the *conditional type* and those in (22a) and (22b) are an example of *quantified type* items. In (21a) and (22a), the numeral is embedded in a UE environment (the second conjunct of a coordinated structure, and the scope of a universally quantified Noun Phrase, respectively); they therefore constitute the UE condition. In

(21b) and (22b), on the other hand, the numeral is embedded in a DE environment (the antecedent of a conditional and a relative clause adjoined to a universally quantified NP, respectively) and thus they constitute the DE condition. Let us note that the first argument of the universal quantifier shares the same semantic property (i.e. DEness) with the antecedent of conditional clauses, which we already discussed in (3). That is, sentence (23a) entails sentence (23b), with the latter considering a subset of the former (i.e. 'the set of guys who drink a *stout beer*' is included in 'the set of guys who drink a *beer*').

(23) a. Every guy who drinks a beer will get an headache
b. Every guy who drinks a stout beer will get an headache

Our predictions are as follows. If numeral strengthening occurs more readily in UE context, people should select the 'exactly' interpretation significantly more often in phrases like (21a) and (22a) than in phrases like (21b) and (22b). Further, if this result is caused by the semantic property affecting numerals interpretation, rather than a specific grammatical construction, we should observe the same trend for both *conditional* and *quantified* sentences.

Participants were asked to make a choice between two alternatives by ticking the one preferred and turn over the page without altering their previous choice. The pivotal question was always posed in the following way: “we are talking about... exactly two cars/at least two cars”. Participants were asked to carry out the task without lingering too much and to answer freely and naturally. It’s worth underscoring once more that the material was almost the same across the experimental conditions. *UE* and *DE* items varied only from two words: the presence of “se” or “che” (“if” and “who” in English) in the latter condition versus the presence of “e” (the conjunction “and”) in the former one.

The data we are focusing on is the percentage of strengthened choices, i.e. the proportion of “exactly N” answers over the totality of answers. The mean strengthened choices percentage (see Tab. 1) for the *conditional type* items was 78% in the *UE condition* vs. 49% in the *DE* condition. The mean for the *quantified type* items was 55% in the in the *UE condition* vs. 27% in the *DE* condition. The *polarity* factor (i.e. UE vs. DE) turned out to be significant, as well as the *type* of item (conditional vs. quantified), whereas no significant interaction was found.

Polarity	Sentence Type	
	Conditional	Quantified
UE	78%	55%
DE	49%	27%

Tab. 1: means of "exactly N" choices

These results clearly show that the linguistic context, more specifically the polarity of the context embedding the numeral determiner, affects the participants’ interpretation choices. Subjects selected significantly more often a strengthened, upper bounded reading in UE contexts with respect to DE ones. This confirms what was found for other scalar terms, like disjunction (Noveck, 2001; Noveck, Chevaux, Guelminger, Sylvestre & Chierchia, 2002; Chierchia, Crain, Guasti, Gualmini, Meroni, 2001) and suggests that the strengthening of numerals is an analogous phenomenon. However, even though a strikingly similar pattern was found in the *conditional* vs. the *quantified* types, the impact of these two types was also a significant factor, as we can see by

looking at the interpretation percentage difference between the *UE* and the *DE* condition (78% - 49% = 29% vs. 55% - 27% = 28%). Accordingly, it must be acknowledged that there are contextual (e.g. the plausibility of the sentence) and structural (e.g. the grammatical construction of the sentence) factors other than polarity that can influence the participants' off-line interpretation of numerals. The influence of polarity is always combined with such factors, which is why the differential behaviour within the same items is important and telling. The crucial point, however, is that the *polarity* and the *type* of items affect participants' choice independently. That is, the entailment property of the context containing the numeral has the same influence on the task regardless of whether the numeral is in a *conditional* or *quantified* type sentence.

5. Experiment 2: an online processing experiment

In Experiment 2, we measured what happens in real time when readers are presented with numerals embedded in UE vs. DE contexts by recording their eye movements. We explored two possible effects. The first is the conventional expectation that if readers commit themselves to one interpretation of a numeral when they read the clause containing the numeral, then reading of the following clause will be disrupted if the interpretation of the numeral is inconsistent with it. The second possible effect is suggested by the proposal advanced in our discussion of the SI hypothesis, that the 'at least' interpretation is the core interpretation and the 'exactly' interpretation is created as a scalar implicature, which may take processing resources. The basic design is to have numerals in UE or DE contexts followed by continuations that do or do not force the upper bounded ('exactly') reading. Reading time on the phrase containing the numeral could reflect any processing cost of constructing an implicature, as explained below. Reading time on the continuation could reflect the cost of revising the initial interpretation of the numeral, and thus provide information about what the initial interpretation is in different contexts.

We prepared a set of two-clause discourses, each containing the first clause of one of the 24 items used in the offline experiment, followed by one of three second clauses, as illustrated in Table 1. The first clause sets up either an upward or a downward entailing context for the numeral, just as in Experiment 1. The second clause could be one of three types: (a) neutral (the same sentences used in Experiment 1), without mention of the entity that was quantified in the first clause; (b) biased towards an upper bounded construal of the numeral in the first clause; or (c) a negative version of the biased continuation. Because of the role of negation, this third continuation canceled the upper bounded reading of the numeral in the first clause, making it functionally equivalent to the neutral reading. The example in Table 1 illustrates.

First line	Second line
(24a) *UE* Giovanni ha **due** macchine in garage e John has **two** cars in the garage and	(25a) *neutral continuation* parcheggia una **motocicletta** nel cortile esterno. he parks a **motorcycle** in the courtyard.
(24b) *DE* Se Giovanni ha **due** macchine in garage If John has **two** cars in the garage	(25b*) positive continuation* parcheggia una **terza macchina** nel cortile esterno. he parks a **third car** in the courtyard. (25c) *negative continuation*

	non parcheggia una **terza macchina** nel cortile esterno. he **doesn't** park a **third car** in the courtyard.
(26a) *UE* Nel mio quartiere ogni ragazza ha **due** fratelli più grandi e In my neighboorhood every girl has **two** older brothers and	(27a) *neutral continuation* desidera una **sorellina** di età inferiore. wishes a younger **sister**.
(26b) *UE* Nel mio quartiere ogni ragazza che ha **due** fratelli più grandi In my neighboorhood every girl who has **two** older brothers	(27b) *positive continuation* desidera un **terzo fratello** di età inferiore. wishes a younger **third brother**.
	(27c) *negative continuation* non desidera un **terzo fratello** di età inferiore. doesn't wish a younger **third brother**.

Tab. 2: Example of experimental items with the different type of continuations

The expected results may be divided into two categories. The first involves the first-pass reading times of the first line. The aim of the first-pass first line analysis is to look for any reading difference between the UE vs. DE contexts, regarding specifically the *numeral* region. If upper bounded ('exactly') readings are preferentially computed in UE contexts, and if they are enrichments of basic lower bounded readings, one might expect slower reading times for the numeral in the UE than in the DE context. The second category of predictions regards the first-pass indices computed on the second line after the *ordinal* (*third car*) is read, and the second-pass indices in all the regions. The key prediction is that only the *positive continuation* with the ordinal forces the upper bounded reading of the numeral in the first sentence. The *positive continuation* is incompatible with the lower bounded reading of the numeral whereas the *neutral* and the *negative continuations* are compatible with such reading. To see this consider the typical positive continuation of the DE conditional sentences, an example of which is repeated here in (28) for convenience.

(28) If John has two cars in the garage, he will park a third car in the courtyard.

If *two* in (28) is not (yet) upper bounded at the level of the antecedent of the conditional, its truth conditional import may be spelled out as follows:

(29a) In any situation in which John has two or more cars in the garage, he will park a third car in the courtyard.

This cannot be true. For take any situation *s* in which John has three cars in the garage; under normal assumptions on ordinals (i.e. assuming that the ordering of cars matches the order of presentation in discourse), a third car will already be in the garage and hence cannot be parked elsewhere. Technically, we have a presupposition clash. The ordinal numeral *third* presupposes that its referent is the third in the most salient ordering available in the context. If John has three cars in the garage, such a presupposition could never be met. Hence this sentence is incoherent (and the same

holds, mutatis mutandis of all other examples of this form). On the other hand, if *two* in (28) is upper bounded, the result is coherent, as the following paraphrase makes clear:

(29b) In any situation in which John has exactly two cars in the garage, he parks a t hird in the courtyard.

So the continuation in (28) does force an upper bounded reading to be embedded in the antecedent of the conditional. Consider next what happens in the *DE* condition when we have the *neutral continuation.*

(30) If John has two cars in the garage, he will park a motorcycle in the courtyard.

Clearly, the continuation in (30) does not conflict with the *at least* interpretation of the numeral, as we may see by considering the following explicit paraphrase:

(31) In any situation in which John has two or more cars in the garage, he will park a motorcycle in the courtyard.

Trivially, the consequent of (30) can be true in any situation *s* in which John parks two or more cars in the garage. Therefore if the numeral gets an *at least* interpretation it will not need to be strengthened since the continuation in (30), unlike the one in (28), does not lead to a contradiction.

Let us now finally turn to a DE sentence followed by a *negative* continuation.

(32) If John has two cars in the garage, he won't park a third car in the courtyard.

In spite of differing minimally from (28), just by the presence of negation, this sentence is not contradictory under the lower bounded construal of the numeral. The following explicit paraphrase may make this claim clear:

(33) In any situation in which John has two or more cars in the garage, he doesn't park a third car in the courtyard.

If John has parked three cars in the garage, the consequent will be automatically (i.e. trivially) true. For (again under standard assumptions on the interpretation of ordinals) it will be true in such an *s* that a/the third car is not parked in the courtyard. If, on the other hand, we are in a situation *s'* in which John has parked only two cars in the garage, then the consequent is consistent with *s'* and hence the conditional in (33) can well be true (and hence non trivial). Again this mode of reasoning applies to all sentences of this shape. Since the only non trivial interpretation of (32) involves the lower bounded reading, we would expect an upper bounded interpretation not to arise at all. Thus, we predict (32) to be *easier* than (28), in spite of the presence of negation, which is usually considered to make things harder.

54 native Italian speakers (between the age of 19 and 29 years old) took part to this experiment. They were told to read the sentences appearing on a screen silently and naturally, in order to answer simple comprehension questions which would follow randomly after reading some sentences. While they read each sentence their eye movements were recorded by an eyetracker.

The 24 experimental items were interspersed with 80 filler sentences and 44 simple comprehension questions. The two experimental factors (*polarity* and *type of continuation*) gave rise to six different conditions (24a+25a, 24a+25b, 24a+25c, 24b+25a, 24b+25b, 24b+25c are examples of the conditions for *conditional items*; 26a+27a, 26a+27b, 26a+27c, 26b+27a, 26b+27b, 26b+27c are examples of the conditions for *quantified items*). Notice that 12 items displayed a conditional clause in the *DE* conditions while the other 12 items contained a universal quantifier restriction in the same conditions, exactly like the first experiment. Thus, items can be divided into several regions of interest, and they were designed so that the regions of major interest (namely the first line numeral and the second line ordinal) were composed by only one word and were the same in every experimental condition (except for the *neutral* continuations where the ordinal was absent). In each item the first clause was ended by a line break, and the second clause (including the conjunction in the UE conditions and the negation in the negative conditions) appeared on the second line.

5.1.1 Results

The first-pass reading indices reveal whether participants spent more time in reading a specific region without having accessed to the following verbal material (i.e. the second clause of the sentence). In the first line analysis (24a and 26a vs. 24b and 26b) only the *polarity* factor was considered as the reading of the continuation did not affect the first-pass reading times of first clause nor did the verbal material in the first clause vary with respect to the *type of continuation* factor. The main point of interest, here, is to check whether the polarity of the context influenced the reading time of the *numeral* region (the word "two", in boldface in 24 and 26). We found a significant main effect of the *polarity* on this region in the *conditioned regression-path duration*. This index equals the *gaze duration time* (the mean of the sum of all fixation times starting with the reader's first fixation inside the region until the reader's gaze leaves the region either to the right or to the left) plus the time spent re-reading just the preceding word, which was the *verb1* region. According to this index participants spent 14ms more in reading the numeral in the UE conditions than in the DE ones, with no significant impact of the type of construction (conditionals vs. quantified sentences) that we employed in this experiment (i.e. the factor *type of item* did not interact with the *polarity*). To ensure that this effect was not caused by a difference in the probability of skipping the *numeral* region we checked that the skipping rate did not vary significantly across DE and UE trials (40.8% and 37.6% of probability of skipping the numeral, respectively).

As for the first-pass reading times on the second line (25a-c and 27a-c), the core results involve the interaction between *polarity* and the *type of continuation* factors. In the *regression-path duration* (the mean of the sum of all fixation times starting with the reader's first fixation inside the region until the reader's gaze leaves the region to the right) computed on the last region (the last words of the sentence, e.g. "in the courtyard") we found significant interactions between *polarity* and the *type of continuation* factors. The interaction for the *neutral* vs. *negative continuation* comparison, in contrast, was not significant. This results show that participants spent more time in re-reading part of the sentence in the DE condition than the UE one, when they encountered the *positive continuation*, whereas this pattern reverted when they read the *neutral* and *negative continuation*. This motivates us to look carefully at the second-pass indices to investigate where the regressions coming from the second line were directed to.

The second-pass indices we will focus our attention on are *second-pass fixation number* (the count of every second pass fixation made on a word) and *second-pass fixation probability* (the probability that the reader made at least a second pass fixation on that word)[3]. In the second-pass indices we found that the *numeral* region, in the first line (24 and 26), displayed the very same pattern of the second line first-pass indices. That is, in the *positive continuation*, in all the second pass indices readers made more regressions towards the *numeral* in the DE conditions than in the UE ones, as displayed in Tab. 2 (second-pass fixation probability: DE: 26% vs. UE 33%; second-pass fixation number: DE: 0.34 vs. UE: 0.46). The *neutral* and *negative continuation* conditions, instead, both displayed the same pattern. Here participants behaved in the opposite way with respect to the *positive continuation*. That is, they made more regressive eye movements towards the *numeral* in the UE conditions than in the DE ones (*neutral continuation.* Second-pass fixation probability: DE: 31% vs. UE 23%; second-pass fixation number: DE: 0.4 vs. UE: 0.3; *negative continuation*. Second-pass fixation probability: DE: 33% vs. UE 25%; second-pass fixation number: DE: 0.46 vs. UE: 0.34). This pattern of results gave rise to significant interactions between the *polarity* and *type of continuation* factors in the *postive* vs. *neutral continuation* and *positive* vs. *negative continuation* comparison, but no interaction between those factors in the *neutral* vs. *negative continuation* comparison.

Index	Continuation			
	Polarity	Positive	Neutral	Negative
Second Pass Fixation Probability	UE	26%	31%	33%
	DE	33%	23%	25%
Second Pass Fixation Number	UE	0.34	0.4	0.46
	DE	0.46	0.3	0.34
Conditioned Sec. Pass Fixation Probability	UE	11%	17%	17%
	DE	16%	13%	9%
Conditioned Sec. Pass Fixation Number	UE	0.14	0.18	0.22
	DE	0.2	0.17	0.12

Tab. 3: second-pass indices for the *numeral* region.

5.1.2 Discussion

The first important finding we have to underscore is that the phrases for which the first experiment participants preferred an upper bounded reading for the numeral in an UE context, exhibit an early processing penalty on the *numeral* region in the second experiment. This effect cannot merely be explained as a general influence of a specific grammatical construction since we tested two different environments (conditional and quantifiers) and the result remained stable across both constructions. It seems that the polarity of the context is a factor systematically exploited by a reader: if the embedding

[3] For a more exstensive analysis of the results of this experiment, see Panizza, Chierchia and Clifton (in press)

context is UE, the upper bounded meaning (on some approaches the SI) is computed (or at least considered) on line, as soon as possible. If the local embedding context is DE, the upper bounded meaning is not considered and if needed, a SI is computed only later while the reader fixates on other regions.

Consider next the effect of the second clause. Our materials were designed to force the *exact* numeral interpretation within the DE context of the *positive continuation,* so that the ordinal numeral was supposed to act as a trigger of a reanalysis if the *exactly* interpretation had not been computed. On the other hand, in the two control sentences this was predicted not to happen. The first control sentence, i.e. the *neutral continuation,* lacked the ordinal numeral, which triggers the need to recalculate the meaning of the numeral in the DE condition. The second, the *negative continuation,* differed by one word from our test sentences (namely, negation), and yet, for semantic reasons, it was not expected to force a reinterpretation of the numeral in the DE condition, in spite of the presence of the ordinal numeral in the second clause. Hence, only the *positive continuation* was expected to force an interpretation of the numeral in the direction of the upper bounded reading. In other words, we expected an interaction between the *polarity* of the first clause and the *type of continuation*.

Our findings are as follows. We found the expected interaction effects in the second-pass indices, resulting in significant interactions between *polarity* and *type of continuation*. The same interaction, along the same direction, was found in the *regression-path duration* computed on the last region, which includes all the regressions made by the reader after reading the whole sentence for the first time. According to these measures, the pattern surfacing from the difference between the *UE* and the *DE* condition of the *negative continuation* is strikingly similar to that coming from the difference between the same conditions of the *neutral continuation*, whereas the pattern in the *positive continuation* is diametrically opposed. Participants made more frequent second-pass fixations when the numeral was embedded under a DE context, in the *positive continuation,* while they behaved in the opposite way in both the *neutral* and the *negative continuations*.

Now, if we compare the first-pass results to those coming from the second-pass we see how in the *positive continuation*, and only there, the participants' reading pattern of the *numeral* in the first line was reversed. During the first-pass it was more difficult to read in the *UE* condition, while according to the second-pass it received more and more often fixations in the *DE* one. In contrast, the readers' behaviour was uniform, across the first and the second-pass, in both the *neutral* and the *negative* continuation. In the latter cases, the numeral always received more first- and second-pass processing in the *UE* condition. This global picture perfectly fits our main claim and its implications for processing. Moreover, to the extent that SIs triggered by other scalar items (like *some* or *or*) are subject to the same generalization, our results provide evidence in favour of the view that the alternation between the lower bounded vs. upper bounded construals of numerals may well be a scalar implicature. In an upward entailing environment a scalar operator is mostly strengthened locally whereas in a downward entailing environment it is typically strengthened only globally. Therefore, in the latter case, additional reading time on the numeral is observed only after the reading of the sentence.

Finally, both early and second-pass effects on the *numeral* region (as well as on the others in the *positive* vs. *negative continuation* comparison) were unaffected by the phrasal structures selected to create a downward entailing environment (*type of item* factor). This shows that the readers' behaviour with respect to the *numeral* region was influenced by the semantic diversity of the two environments (DE vs. UE), in

interaction with the type of continuation, rather than other contingent factors like the specific words or the syntactic construction adopted in building in the sentences.

6. Conclusions

The results of these experiments show that structural factors (entailment properties of the local context) affect the interpretation of numerals. Furthermore they bring evidence that the lower bounded interpretation of numerals occurs preferentially in DE contexts (with respect to minimally different UE ones), while the upper bounded one occurs preferentially in UE contexts (with respect to minimally different DE ones).

We have addressed this task by investigating two types of functors, *every* and *if* (both DE in their restriction) and contrasting them with minimally different UE contexts (*and*, in the case of *if*; while in the case of *every*, we have simply displaced the numeral from the restriction to the scope). The choice of material (and the various controls we have run) makes it implausible that our results may stem from idiosyncratic features of the selected items or of the context. An off-line questionnaire confirmed that readers interpret numerals in conformity with (4) most of the time. An on-line experiment based on the recording of eye movements seems to reveal a systematic processing penalty associated with contexts in which one forces readings that go against the generalization (4).

These findings have rich consequences. For one thing, they support something like 'Optimize Informativeness', i.e. the idea that the readings of numerals are unconsciously chosen in such a way as to avoid interpretations that lead to weakening with respect to available alternatives (unless forced to). This suggests that the parser somehow checks entailments in selecting a reading. Of great interest, for the future research, might be to investigate how such principle interacts with the contextual factors (i.e. the saliency of an entailing scale in a given context). A plausible hypothesis, stemming from this work, buys on the idea that the kind of computations performed while drawing a scalar implicature are encapsulated from extra-linguistics factors. This is not tantamount to saying that extra-linguistics factors do not affect the final outcome of scalar processes, but rather that the scalar computation itself is performed by a cognitive system relatively blind to contextual information. Where the context may be playing a crucial role is in determining whether the scalar alternatives are active or not. Although this 'blindness to context' hypothesis remains still rather speculative, the lack of interaction between the effects of polarity and type of constructions that we used in the experimental items could be taken as going in this direction.

A second important point that emerges is the following. Other scalar terms (e.g. quantifiers like *some*, connectives like *or*, etc.) have been argued to be subject to a similar constraint in their interpretation. This suggests that the variation in meaning of numerals is probably due to one and the same mechanism, presumably a scalar implicature, no matter how much numerals may otherwise differ in meaning and processing from, e.g., quantifiers like *some* (cf. on this Huang & Snedeker, 2009). More experimental work is called in to test whether other scalar dimensions are affected at the same way by the experimental manipulation of the polarity of the embedding context.

The points made above are quite general. There are many other elements emerging from the present study that we think are relevant to ongoing theoretical debates on the nature of implicatures. One worth mentioning is the following. With

numerals it is easy to see (and, as a matter of fact, quite uncontroversial) that both lower and upper bounded readings can occur in embedded positions. If the distribution of these readings follows a pattern similar to that of other scalar terms, and hence may be due to the same general mechanism responsible for SIs, we would have further confirmation of the existence of embedded implicatures, a frequently disputed claim (defended by Chierchia et al., in press). Be that as it may, we think the present work provides strong evidence in favour of the fact that the polarity of the context where the numeral is embedded affects both the interpretation and the processing of of numerals.

References

Atlas, J., & Levinson, S. C. (1981). It-clefts, informativeness, and logical form: Radical pragmatics (revised standard version) in P. Cole (Ed.) *Radical Pragmatics* (pp. 1-61). New York: Academic Press.

Breheny, R., Katsos, N., & Williams, J. (2006). Are Scalar Implicatures Generated by Default? *Cognition* 100(3), 434-463

Breheny, R. (2008). A *New Look* at the Semantics and Pragmatics of Numerically Quantified Noun Phrases. *Journal of Semantics* 25(2), 93-139.

Carston, R. (1998) "Informativeness, relevance and scalar implicature", in R. Carston and S. Uchida (Eds.), Relevance Theory: applications and implications, (pp. 179-236). Amsterdam: Benjamins.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond*. Oxford: Oxford University Press.

Chierchia, G. (2006). Broaden your Views. Implicatures of Domain Widening and the "Logicality" of Language. *Linguistic Inquiry*, 37(4), 535-590.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: evidence for a grammatical view of scalar implicatures. In A. H.-J. Do, L. Dominguez, & A. Johansen (Eds.), *Proceedings of the 25th annual Boston University conference on language development*, (pp. 157–168). Somerville, MA: Cascadilla Press.

Chierchia, G., Spector, B., & Fox, D. (in press). The Grammatical View of Scalar Implicatures and the Relationship between Semantics and Pragmatics, to appear in Portner, P., Maienborn, C., & von Heusinger, K. (Eds.), *Handbook of Semantics*, Mouton de Gruyter.

Chierchia, G., Frazier, L., & Clifton, C. Jr. (in press). When basic meanings are (not) enough: Processing scalar implicatures in adult comprehension. In M. Biezma, C. Davis, and J. Harris, (Eds.), *University of Massachusetts Occasional Publications in Linguistics,* (Vol. 39) (Issues in Pragmatics).

Grice, H. P. (1967a), Logic and Conversation. In P. Cole and J. L. Morgan (Eds.),

Syntax and Semantics, vol. 3: Speech Acts (pp. 41-58). New York: Academic Press, 1975. Reprinted in H. P. Grice 1989, (pp. 22-40).

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

Horn, L. (1972). On the semantic properties of the logical operators in English. Doctoral dissertation, UCLA.

Horn, L. (1992). The said and the unsaid. *Proceedings of SALT* 2 (pp. 163-191). Department of Linguistics, Ohio State University.

Horn, L. R. (2004). Implicature. In L. R. Horn and G. Ward (Eds.), *The Handbook of Pragmatics.* Malden, MA: Blackwell.

Horn, L. R. (2007). Issues in the investigation of implicature. To appear in Petrus, K. (Ed.), *Volume in honor of Grice*.

Horn, L. R.: (1989). *A Natural History of Negation*. Chicago, Ill.: University of Chicago Press.

Huang, Y. & Snedeker, J. (2009). On-line Interpretation of Scalar Quantifiers: Insight into the Semantic-Pragmatics Interface. *Cognitive Psychology*.

Kadmon, N. (2001). *Formal Pragmatics*. Cambridge, Mass.: Blackwell.

Landman, F. (2003). Predicate-argument mismatches and the adjectival theory of indefinites. In M. Coene & Y. D'hulst (Eds.), The Syntax and Semantics of Noun Phrases (pp. 211–237). *Linguistics Today 55*. John Benjamins: Amsterdam and Philadelphia.

Levinson, S. (2000). *Presumptive meanings*. Cambridge, MA: MIT Press.

Musolino, J. (2004). The semantics and acquisition of number words: integrating linguistic and developmental perspectives, *Cognition* 93(1), 1-41.

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition,* 78(2), 165-188.

Noveck, I. A., Chierchia, G., Chevaux, F., Guelminger, R., & Sylvestre, E. (2002). Linguistic-pragmatic factors in interpreting disjunctions. *Thinking and Reasoning,* 8(4), 297-326.

Panizza, D., Chierchia, G. & Clifton, C. (in press). On the role of entailment patterns and scalar implicatures in the processing of numerals. *Journal of Memory and Language*.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition*, 86(3), 253-282.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin,* 124, 372-422.

Rooth, M. (1985). *Association with Focus*. University of Massachusetts at Amherst Ph. D. dissertation.

Sperber, D., & Wilson, D. (1985/1995). *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.

Name and Affiliation

Daniele Panizza, Department of Cognitive Science, University of Trento, Rovereto.
Gennaro Chierchia, Department of Linguistics, Harvard University, Cambridge, MA.

Keywords

Numerals, implicature, scalar implicature, semantic processing, inference, entailment, polarity, semantics, pragmatics, informativeness