Published in final edited form as:
J Mem Lang. 2009 November 1; 61(4): 503–518. doi:10.1016/j.jml.2009.07.005.

On the role of entailment patterns and scalar implicatures in the processing of numerals

Daniele Panizza[1], **Gennaro Chierchia**[2], and **Charles Clifton Jr.**[3]
[1]Department of Cognitive Science, University of Trento, Rovereto. daniele.panizza@unitn.it
[2]Department of Linguistics, Harvard University, Cambridge, MA. chierch@fas.harvard.edu
[3]Department of Psychology, University of Massachusetts, Amherst, MA. cec@psych.umass.edu

Abstract

There has been much debate, in both the linguistics and the psycholinguistics literature, concerning numbers and the interpretation of number denoting determiners ('numerals'). Such debate concerns, in particular, the nature and distribution of upper-bounded ('at-least') interpretations vs. lower-bounded ('exact') construals. In the present paper we show that the interpretation and processing of numerals are affected by the entailment properties of the context in which they occur. Experiment 1 established off-line preferences using a questionnaire. Experiment 2 investigated the processing issue through an eye tracking experiment using a silent reading task. Our results show that the upper-bounded interpretation of numerals occurs more often in an upward entailing context than in a downward entailing context. Reading times of the numeral itself were longer when it was embedded in an upward entailing context than when it was not, indicating that processing resources were required when the context triggered an upper-bounded interpretation. However, reading of a following context that required an upper-bounded interpretation triggered more regressions towards the numeral when it had occurred in a downward entailing context than in an upward entailing one. Such findings show that speakers' interpretation and processing of numerals is systematically affected by the polarity of the sentence in which they occur, and support the hypothesis that the upper-bounded interpretation of numerals is due to a scalar implicature.

Keywords

numbers; numeral determiners; scalar implicature; polarity; entailment; semantic processing

---

Numerical expressions (‘numerals’) are known to have (minimally) two interpretations or construals: an ‘exact’ (or upper bounded) reading and an ‘at-least’ (or lower bounded) reading (the upper/lower bounded terminology is standard; Horn, 1972). They are exemplified by (1a) and (1b) respectively (and similarly for (2a) and (2b)):

```
(1)    a. If I find my wallet, I’ll lend you ten dollars.
```


Address correspondence to: Daniele Panizza, Dipartimento di Scienze della Cognizione e della Formazione, Corso Bettini n. 31, 38068 Rovereto (TN), Italy. Telephone: 0039-0464-808668, Fax Number: 0039-0464-808655, daniele.panizza@unitn.it.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

```
      b. If I lend you ten dollars, I'll never see my money again.

(2)   a. Everyone will lend John ten dollars.

      b. Everyone who lends John ten dollars will never see his money again.
```

The most salient reading of (1a) is an upper bounded one (…, I'll give you exactly ten dollars), while (1b) is more naturally interpreted as lower bounded (If I lend you ten dollars or more, …). Similarly, mutatis mutandis, for the pair of sentences in (2).

In the present paper, we address three questions:

- **a.** Do speakers consistently and generally show the differences in numeral interpretive preferences we see in (1) and (2)?
- **b.** Can such a preference be attributed to structural factors, in particular the entailment patterns (see below) associated with the contexts?
- **c.** Is there a processing cost to arriving at one interpretation compared to the other?

We address these questions in two experiments. The first is a simple questionnaire designed to determine whether the difference between contexts like (1a, 2a) and (1b, 2b) systematically affects interpretations, while holding non-structural factors as constant as possible. We anticipate that 'exact' interpretations will be relatively preferred in contexts like (1a) and (2a), and 'at-least' interpretations in contexts like (1b) and (2b). The second is an eyetracking study in which we concentrate on two phenomena. We first concentrate on the difficulty of reading a numeral in contexts like (1a) and (2a) vs. (1b) and (2b), in an attempt to identify any processing costs associated with arriving at the interpretation preferred in one context. We anticipate seeing increased processing costs when the 'exact' interpretation is encouraged, for reasons we detail below. We then concentrate on the difficulty of reading a following clause that is consistent with the 'exact' interpretations of the numeral in the first clause (compared with clauses that are consistent with an 'at-least' interpretation), on the assumption that difficulty will be reduced if the 'exact' interpretation had been made in the first place, in particular, following contexts like (1a) and (2a).

The linguistic background: Numerals and entailment

We first discuss the semantics of numerals, and then discuss what it is about the contexts in (1) and (2) that results in the apparent interpretive differences. There is general agreement, in the current linguistics literature, that numerals can have an 'exact' and an 'at-least' interpretation. The disagreement concerns how they come to get such interpretations. On the basis of the background in the literature, the main theoretical choices can be given in a compact form shown in Table 1. Symmetric theories are typically 'lexicalist' in nature, in the sense that ambiguity or underspecification is located in the lexical items. Asymmetric theories claim that the lexical entries have a basic meaning, and the alternate construal is derived by compositional semantics or by some post-compositional, pragmatic process.

One way to evaluate these different views of numeral semantics is to examine how they account for the differing effects of contexts like (1) and (2). Following Ladusaw (1979), Chierchia (2004), and Chierchia, Fox, and Spector (2008), we take the critical difference between contexts like (1a) and (2a), on the one hand, and (1b) and (2b), on the other, to be their entailment patterns. An entailment relation is a semantic (and logical) property that holds between two or more propositions. For instance, sentence (3a) entails (3b) because in every situation where (3a) is true, (3b) is true as well.

```
(3)    a. I ate pizza with anchovies.

       b. I ate pizza.
```

The entailment pattern associated with affirmative sentences, such as in (3) (and in (1a) and (2a)), is *upward entailing*. An upward entailing context is one that licenses inferences from subsets to supersets. The sentence in (3a), which is a statement about a subset, entails (3b), a statement about a superset (with the denotation of 'pizza with anchovies' being a subset of the denotation of 'pizza'). Similarly, a clause in the consequent of a conditional, as in (4), and a clause in the second argument (the scope) of a universal quantifier like *every*, as in (5), are upward entailing. In both cases, the subset (4a, 5a) entails the superset (4b, 5b).

```
(4)    a. If I go home, I'll eat pizza with anchovies.

       b. If I go home, I'll eat pizza.

(5)    a. Every boy that goes home will eat pizza with anchovies.

       b. Every boy that goes home will eat pizza.
```

However, the antecedent of a conditional, as in (6), and the first argument of a universal quantifier (the restriction), as in (7), is a *downward entailing* context. Downward entailing contexts invert the entailing pattern of upward entailing contexts, that is, they licence inferences from supersets to subsets (e.g. 6a and 7a entail 6b and 7b, respectively).

```
(6)    a. If I eat pizza, I'll get sick.

       b. If I eat pizza with anchovies, I'll get sick.

(7)    a. Every boy that ate pizza got sick.

       b. Every boy that ate pizza with anchovies got sick.
```

Entailment patterns are most closely associated with polarity phenomena: negative polarity items like *any* and *ever* occur only in downward entailing contexts (Ladusaw, 1979). Entailment patterns have also been claimed to be associated with scalar implicatures (Chierchia, 2004; cf. Gazdar, 1979). 'Exact' interpretations of numerals seem to be preferred in upward entailing contexts, like (1a) and (2a); 'at-least' interpretations seem to be relatively preferred in downward entailing contexts. We advance the claim in (8):

```
(8)    The upper bounded interpretation of numerals occurs preferentially in
upward
       entailing contexts with respect to minimally different downward
entailing contexts.
       The lower bounded interpretation of numerals occurs preferentially in
downward
      entailing contexts with respect to minimally different upward entailing
contexts.
```

Notice that (8) does not say anything about the proportion of lower bounded vs. upper bounded interpretations intended by the speaker in any context. Nor does it deny that contextual factors affect the interpretation of a scalar item (cf. Breheny, Katsos, & Williams, 2006; Breheny, 2008). Our thesis posits that, if we try to keep the role of world knowledge and other extralinguistic factors constant, the 'exact' interpretation of a numeral is more likely to emerge in an upward entailing context while the 'at-least' interpretation is more likely to emerge in a downward entailing one.

This claim has processing implications. Although (8) per se does not say anything about the cost of deriving an upper/lower bounded reading, it raises the possibility that one interpretation is basic and arriving at the other is relatively costly. Further it suggests that forcing a sentence to have an upper bounded interpretation of a numeral is more costly if the numeral had occurred in a downward entailing context than if it had occurred in an upward entailing context. Dispreferred interpretations of a word or a sentence impose a processing load, with respect to the preferred one (cf. Rayner, 1998, for a review of research using the eyetracking methodology we will use).

Grant, for the sake of discussion, that (8) holds. Why would it hold? In upward entailing contexts, the 'exact' interpretation of a numeral logically entails the 'at-least' one; classical ('normative') logic tells us that (9a) entails (9b):

```
(9)    Upward entailing: exact → at-least

       a. John owns exactly three cars

       b. John owns at least three cars
```

In downward entailing contexts, this pattern is reversed: the 'at-least' interpretation entails the exact one. For example, (10a) entails (10b)

```
(10)    Downward entailing: at-least → exact

        a. Every one who owns at least three cars is from Boston

        b. Every one who owns exactly three cars is from Boston
```

So the 'exact' interpretation leads to a logically stronger (and hence more informative) statement in an upward entailing interpretation; it leads to a logically weaker (and hence less informative) statement in a downward entailing context. We conjecture that this would be a very good reason for claim (8) to hold. When given a choice, speakers (and hearers) prefer the interpretation that maximizes the informative content of their utterances (or, if you wish, they behave as if they had such a preference).

Establishing (8) and investigating the processing load associated with different construals of numerals have important theoretical consequences for understanding the semantics of number words and for understanding how scalar words in general work. In the following subsection we try to briefly explain why.

Scales and scalar implicatures

Numerals are scalar terms, where the scale is the number line. They are likely to share properties with other scalar terms, including quantifiers (where the scale is *some-all*), connectives (*or-and*), modals (*may-must*), scalar adjectives (*cold-warm-hot*), etc. (cf. Horn, 1972). The literature contains claims similar to (8) about scalar terms in general, particularly in the context of work on scalar implicatures (Grice, 1967,1989). Consider the quantifier *some*. In these analyses, the ‘some and possibly all’ meaning of *some* is taken to be basic. The other interpretation of *some*, ‘some but not all,’ is a ‘stronger’ interpretation, in the sense that it is more restrictive: there are situations in which the stronger interpretation is false but the basic interpretation is true. The stronger interpretation is taken to be an implicature (a ‘scalar implicature,’ sometimes referred to as ‘pragmatic strengthening’), and the question at issue is, why and when are scalar implicatures made.

Contrast the use of *some* and *two* in a upward entailing context (11a) and in a downward entailing context (11b):

```
(11)    a. John argued with some/two of his teachers, and he’ll be summoned
by the principal.

        b. If John argued with some/two of his teachers he’ll be summoned by
the principal.
```

In sentence (11a) *some* seems to mean ‘some though not all’ (upper bounded interpretation), while clearly that isn’t how *some* is interpreted in (11b) (which displays a lower bounded interpretation, compatible with John’s arguing with all of the teachers). Likewise, *two* in (11a) seems to mean 'two though not more' whereas in (11b) it's more compatible with the lower bounded meaning ('two and maybe more'). Grammatical contexts like these can be shown to have similar effects on the other scalar terms mentioned above.

The asymmetric proposals outlined in Table 1 provide an illuminating account of these effects. Consider first the Gricean and the Neogricean analyses of this observation (e.g. Levinson, 2000; Horn, 1972, 2007). A reader or listener is assumed to pragmatically enrich the interpretation of a sentence by following some simple assumptions about how conversation proceeds – the Gricean maxims, including the maxim of informativeness (i.e. quantity). Consider (12a), which is logically stronger or more informative than (11a) (the logically more informative alternative entails the less informative one, just as the statement about a subset *I ate pizza with anchovies* in (3a) entails the less informative superset statement *I ate pizza*, (3b)).

```
(12)    a. John argued with all/three of his teachers, and he’ll be summoned
by the principal.

        b. If John argued with all/three of his teachers he’ll be summoned by
the principal.
```

In a (neo)Gricean account, the reader or listener of (11a) considers the alternatives to what was said – in this case, the relevant alternative is (12a) – and realizes that this alternative is logically stronger or more informative than (11a). S/he reasons, if the writer/speaker had reason to believe the more informative proposition, s/he would have produced it. Therefore, on the additional assumption that the writer/speaker is well-informed, the Gricean reasoner concludes that *some* in (11a) implicates *some but not all* as well as *two* implicates *two but not three (or* *more)*. This reasoning results in a strengthened statement: the implicature *John argued with some/two but not all/more of the teachers* is logically stronger than the basic interpretation *John argued with some/two and possibly all/more of the teachers*. (Note that this is an implicature, not an implication. It can be cancelled. One can say *John argued with two of his teachers; in fact, he argued with six of them*).

However, this logic does not go through in a downward entailing context, like (11b). In a downward entailing context, the relevant entailments involve the whole sentence, and are reversed. The relevant alternative to a sentence that contains *some* (as in 11b) is a sentence that contains *all* (as in 12b). Here, the original assertion (11b) is stronger and more informative, and entails its alternative (12b). If arguing with some and possibly all of the teachers results in the possibility being summoned by the principal, then arguing with all of the teachers would also do so. A writer or speaker of (11b) following the Gricean maxim of informativeness would therefore be committed to (12b), and the reader/listener would not be justified in adding an implicature to (11b) so that it means *If John argued with some/two but not all/more of the teachers, he'll be summoned by the principal.*

These facts suggest that the reader or listener must go through some rather complicated reasoning processes to draw scalar implicatures in the right places, construing a basic meaning (some and possibly all, at least n) as a strengthened (some but not all, exactly n) reading when appropriate. Other theorists (Chierchia, 2004; Chierchia et al., 2008; cf. Gazdar, 1979) have suggested that the logic underlying this reasoning has been incorporated into the grammar, so that strengthening (making a scalar implicature) is permitted in an upward entailing environment but discouraged in a downward entailing environment.

The positions discussed so far assume that the 'at least, some and possibly all' meanings are basic, and the 'exactly, some but not all' meanings are created as implicatures. It is possible to take the opposite position, at least for numerals (Breheny, 2008) according to which the 'exact' meaning is basic. Rather than discussing the details of this position, we will evaluate it in the light of our experimental data. To anticipate, such a position suggests that their might be a processing cost in a context that favors an 'at-least' interpretation. In contrast, the position that says that the basic meaning of numerals is the 'at-least' interpretation suggests that any processing cost will appear in contexts that support the 'exact' interpretation.

Other theorists take a "symmetric" position (Table 1), at least regarding numerals (if not other scalar terms). Some, such as Horn (1972) suggest that numerals are ambiguous between an 'exact' and an 'at-least' reading. However, such a position does not intrinsically account for why scalar terms are sensitive to the entailment context (other ambiguous words, such as *bank*, are not sensitive to such contexts) and it does not account for why the effects of entailment contexts are apparently similar for numerals and for other scalar terms (which are presumably not ambiguous). Similarly, "relevance" theorists take the meaning of numerals (and possibly other scalar terms) to be underspecified, relying on the content of the context in which they are used to specify how they are construed. We suspect that both ambiguity theorists and relevance theorists could incorporate the logic we sketched above for why upward entailing and downward entailing contexts differ in the interpretations that they favor, but it does not seem straightforward for such a theorist to provide an account for why one interpretation should result in greater processing cost than the other.

Some experimental results on scalar implicatures

There is small but growing literature on scalar implicatures, mostly on scales other than numerals, which we will review briefly. There is some evidence favoring a version of our claim (8), about upward entailing vs. downward entailing contexts, applied to scalar terms other than numerals (Chierchia et al., 2001; Chierchia, Frazier, & Clifton, in press; Noveck et al., 2002;

and Schwarz, Clifton, & Frazier, in press). Noveck et al. (2002) used a reasoning task to show that subjects were more likely to take an exclusive (strengthened) reading of *or* in an upward entailing than a downward entailing context. Chierchia et al. (2001) employed the Truth Value Judgment Task (cf. Crain & Thorton, 1998) to show that both adults' and children's interpretation of *or* depended on whether it occurred in a relative clause modifying a universally quantified noun (the first argument of the universal quantifier, which is a downward entailing context) or in the phrase predicated of the universally quantified noun (the second argument, which is upward entailing) (cf. example (2), presented earlier). Both adults and children accepted the strengthened (exclusive or) interpretation more often when *or* had occurred in the second than the first argument. Chierchia et al. (in press) and Schwarz et al. (in press) tested adults with written and auditory questionnaires and self-paced reading, reporting that strengthened interpretations of *some* and *or* were more preferred in upward entailing than in downward entailing contexts.

Some research indicates that drawing scalar implicatures requires linguistic sophistication and processing time. Papafragou and Musolino (2003) showed that children were more likely than adults to accept statements with weaker scalar terms (*some, two, start*) in contexts that supported a stronger term (*all, three, finish*) (cf. Noveck, 2001; note, in Papafragou and Musolino, children were more likely to reject statements with numbers than statements with the other terms). Bott and Noveck (2004; cf. Novek & Posada, 2003) found longer response times when participants judged underinformative statements such as *Some elephants are mammals* to be false, than when they judged them to be true, arguing for a late processing penalty induced by the scalar implicature. Breheny et al. (2006), in an on-line experiment, investigated the processing of a quantifier occurring after an introductory mini-discourse using a self-paced reading task. Their subjects read short discourses like (13)

(13) a. Mary asked John whether he intended to host all his relatives in his tiny apartment. John replied that he intended to host (only) *some of his relatives*. The rest would stay in a nearby hotel.

b. Mary was surprised to see John cleaning his apartment and she asked the reason why. John told her that he intended to host some of his relatives. The rest would stay in a nearby hotel.

The word *all* in the first sentence in (13a) is meant to bias the interpretation of the second sentence towards the upper bounded reading; the first sentence in (13b) is meant to bias the interpretation of *some* towards the lower bounded reading. Including *only* in (13a) creates a control condition that disambiguates the sentence. The objective was to find the impact of this bias on the processing of *some*. The authors reported two informative effects. First, reading times for the trigger region *some of his relatives* were longer in (13a) than in (13b), suggesting that there was a processing cost to making the semantic implicature (that not all of the relatives would be hosted). Second, the region containing *the rest* was read more slowly in (13b) than (13a) (with or without *only*), suggesting that the semantic implicature had not been made in (13b) and the existence of some non-hosted relatives had to be inferred when the last sentence of (13b) is read. Brehehy et al. claim that their results argue against the view that scalar implicatures are computed by default (eg., Levinson, 2000) (but cf. Bezuidenhout & Cutting, 2002, whom Breheny et al. criticize for having inadequate materials).

In a recent visual world eyetracking study, Huang and Snedeker (2009) also found evidence for a cost in making a scalar implicature with *some*. They presented their subjects with arrays of four pictures, including, in the critical trials, one girl with two socks, a boy with one sock, a girl with three soccer balls, and a boy with nothing, and played them a sentence like (14):

```
(14)    Point to the girl that has some of the socks.
```

Note that the critical word, *socks*, has the same onset as the potential distractor, *soccer balls*. In this condition participants were not able to disambiguate the sentence until 800 milliseconds after the onset of the quantifier ("some of the"), in that they did not look significantly more often at the girl with two (i.e. some but not all of the) socks until this time. It appears that the scalar implicature (*the girl has some but not all of the socks*) was delayed, as the participants' looks converged towards the referent almost at the onset of the disambiguation ("socks").

A second condition in the Huang and Snedeker (2009) visual world studies presented listeners with sentences like (15)

```
(15)    Point to the girl that has two of the socks
```

In this case, while looking at a scenario very similar to that for *some*, participants disambiguated the girl with two socks vs. the girl with three soccer balls at about 400 milliseconds since the onset of the word "two". Further, they found that participants disambiguated during the same time window (400–00 ms) the referent of descriptions involving unambiguous quantifiers such as *all* and *three*, in the very same scenario. The key point, here, is that if *two* is initially interpreted with a lower bounded meaning ('at least two') the girl with two of the socks and the one with three soccerballs would both be good referents for the description ("the girl that has two of the ...."). The failure to show a delayed interpretation of *two* (compared to the delay with *some*) led the authors to conclude that numerals have an 'exact' lexical semantics that allows listeners to pick the right referent very quickly, whereas *some* requires some kind of pragmatic work (i.e. the computation of a scalar implicature) over the semantic meaning, which is lower bounded for *some*. This part of Huang and Snedeker's results should be kept in mind as it appears to contrast with the findings we will report. We will return to this issue in our final discussion.

In sum, there is growing evidence that making scalar implicatures is dependent on the context in which a scalar term appears, and when a scalar implicature is made, doing so is costly in processing terms. However, to our knowledge, there is no evidence about the processing costs of making semantic implicatures in upward entailing vs. downward entailing contexts, nor has this claim been investigated with numerals. The experiments we report below test our claim (8) in the domain of numerals, and examine processing costs in upward entailing vs. downward entailing contexts.

The current experiments

Our first experiment is a simple questionnaire, designed to evaluate the accuracy of our claim (8) about the interpretation of sentences with numerals. While intuitions do seem to support this claim, we believe that it is useful to establish that people untrained in linguistics share our intuitions, and that the intuitions hold systematically across different instances of downward entailing and upward entailing contexts. The second experiment is a reading experiment that measures readers' eye movements to infer processing difficulty of sentences based on those used in Experiment 1. It is patterned on the Breheny et al. (2006) Experiment 3 described above, except that it uses the entailment properties of grammatically-defined contexts rather than using contexts in which *some* contrasts with a preceding *all*. It is designed to see whether or not there is a cost to (a) making a scalar implicature when reading a numeral phrase, and (b) interpreting a following clause that presupposes an 'exact' interpretation of a numeral in a downward entailing context, in which we argue a scalar implicature is relatively dispreferred. We take success in showing these processing costs to argue in favor of the asymmetric interpretation of numeral phrases described in Table 1. If the interpretation of numerals is indeed sensitive to whether they occur in a downward entailing vs. upward entailing contexts, and if we find evidence that one reading of numerals is more basic than the other one, then the above theories will be either confirmed or disconfirmed depending on how well they are able to derive this generalization.

Experiment 1

This experiment was designed to determine the relative frequencies of 'exact' and 'at-least' interpretations of numerals in different grammatical contexts and to develop materials for the eye tracking experiment to follow. We explicitly asked our native Italian-speaking participants to indicate their interpretation of the numeral determiner in a written questionnaire. They had to choose between the stronger 'exact' meaning and the weaker 'at-least' one by checking the appropriate box, after reading a sentence of the kinds displayed in (16) and (17).

(16) a. Giovanni ha **due** macchine in garage e parcheggia una motocicletta nel cortile esterno.
John has **two** cars in the garage and he will park a motorcycle in the courtyard.

b. Se Giovanni ha **due** macchine in garage, parcheggia una motocicletta nel cortile esterno.
If John has **two** cars in the garage he will park a motorcycle in the courtyard.

(17) a. 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.

b. 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.

The (a) and (b) forms of the sentences in (16) and (17) differ minimally. Two different sentence forms with different ways of manipulating entailment patterns were used, for generality. Items (16a) and (16b) are *conditional* sentences and those in (17a) and (17b) are *quantified* sentences. In (16a) and (17a), the numeral is embedded in an upward entailing environment (the second conjunct of a coordinated structure, and the scope of a universally quantified noun phrase). In (16b) and (17b) the numeral is embedded in a downward entailing environment (the antecedent of a conditional and a relative clause attached to a universally quantified NP, respectively). If numeral strengthening occurs more readily in a upward entailing context, as we claimed in (8), people should select the ‘exact’ interpretation significantly more often in phrases like (16a) and (17a) than in phrases like (16b) and (17b).

Method

Subjects—The subjects were 50 native speakers of Italian, between the ages of 19 and 26. They were undergraduate students enrolled in a Psychology course at the University of Milan-Bicocca.

Procedure and materials—Twenty-four items, 12 of the *conditional type* and the other 12 of the *quantified type*, were counterbalanced among the two experimental conditions (12 *upward entailing* vs. 12 *downward entailing*) and combined with four unambiguous control items that were included to make sure subjects were answering carefully. All items will be made available on [EDITOR – Link to Elsevier website] They were put together in a paper questionnaire, one item per page, and participants were asked to make a choice between two alternatives by ticking the one preferred and turning the page without altering their previous choice. The pivotal question was always posed (in Italian) in the following way (18):

```
(18)            we are talking about…
        exactly two cars              at least two cars
```

Participants were asked to do the task without lingering too much and to answer freely and naturally.

It’s worth emphasizing that the material was almost the same across the experimental conditions. Upward entailing and downward entailing items differed only in 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. All the other words remained unchanged. Eight different lists were given to the participants: each list contained 6 items like (16a), six like (16b), six like (17a) and six like (17b), and conditions were counterbalanced within subjects (e.g. 1a, 2b, 3a, 4b, and so on, for each type of item).

Results and discussion

Three subjects were excluded from the analysis as they answered some of the control items incorrectly. 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 participants’ choices were analyzed with a Generalized Linear Mixed Model (GLMM) assuming a binomial family distribution and adopting as fixed factors the *polarity* of the phrase (*upward* vs. *downward entailing* conditions) and its *type* (*conditional type vs. quantified type*), whereas we treat items and subjects as random factors (cf. Baayen, Davidson & Bates, 2008; Jaeger, 2008). The mean percentages of the strengthened choices for the *conditional type* items were 78% in the *upward entailing condition* vs. 49% in the *downward entailing* condition. The means for the *quantified type* items were 55% in the *upward entailing condition* vs. 27% in the *downward entailing* condition. GLMM analysis showed a significant main effect of the *polarity* factor (p. <.001) and a significant main effect of the *type* factor (p. < .001). The interaction between these two factors was not significant (p = .454).

These results clearly show that the polarity of the semantic function embedding the numeral determiner affects the participants’ interpretation choices. Subjects selected significantly more often a strengthened, upper bounded reading more often in upward entailing contexts than in downward entailing contexts. This confirms what has been found for other scalar terms, like *or* and *some* (Chierchia et al., 2001; Chierchia, Frazier, & Clifton, in press; Noveck, 2001;

Noveck et al., 2002; and Schwarz, Clifton, & Frazier, in press), and suggests that the strengthening of numerals is an analogous phenomenon.

The mean percentage of strengthened choices was 63.5% for the conditional sentences vs. 41% for the quantified sentences. However, the effect of upward vs. downward entailing context was remarkably similar for both sentence types, 29% vs. 28%. We acknowledge that, besides polarity, there are contextual factors (e.g. the plausibility of the sentence) and structural factors (e.g. the grammatical construction of the sentence) that can influence the participants' off-line interpretation of numerals. 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.

Summing up, in an off-line task, subjects favor upper bounded readings more often when the numeral is embedded within upward entailing contexts than in downward entailing ones. Since the materials in the two contexts were minimally different, it is implausible that things like world knowledge are the deciding factor. This provides evidence in favor of our main claim, (8).

Experiment 2

NIH-PA Author Manuscript

In our second experiment, we measured what happens in real time when readers interpret numerals embedded in upward entailing vs. downward entailing contexts by recording their eye movements. The basic design is to present numerals in upward entailing or downward entailing contexts followed by continuations that force or do not force the upper bounded ('exact') reading. We explored two possible effects. The first is suggested by the proposal advanced in our discussion of the scalar implicature hypothesis, that the 'at-least' interpretation is the core interpretation and the 'exact' interpretation is created as a scalar implicature, which may take processing resources. Reading time on the phrase containing the numeral could reflect any processing cost of constructing a semantic implicature, which we expect to occur in an upward entailing context. The second possible effect 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. Reading time for the continuation could reflect the effect 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 Experiment 1, followed by one of three second clauses, as illustrated in Table 2. The first clause established either an upward or a downward entailing context for the numeral, just as in Experiment 1. The second clause was either (a) neutral (the same sentences used in Experiment 1), not mentioning again 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 negated 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 equivalent to the neutral reading.

NIH-PA Author Manuscript

It is possible to view the positive continuation (labelled '2-B' in Table 2) as the 'experimental' continuation (in that it should cause difficulty for a reader who has not assigned an exact numerical interpretation). Its reading time can be compared against each of the two control conditions (2-A and 2-C in Table 2). Neither of these control conditions requires the 'exact' interpretation of the first clause, so should not cause difficulty when that interpretation has not been made. The neutral continuation (2-A) has the advantage of being very natural and easy to understand, but does introduce different lexical items than are used in the experimental (2-

B) continuation (which is not a serious shortcoming, since we looked for an interaction between the first and second clause forms). The negative continuation (2-C), on the other hand, contains the same lexical material as the positive (experimental) continuation with the exception of a negative particle (*non*, which is *not* in English) at the beginning of the second clause. We explain below why the negative continuation does not presuppose a strengthened 'exact' interpretation of the numeral, but we acknowledge that prior to the experiment it was not certain that our subjects would be fully sensitive to the considerations behind this explanation.

The predicted results may be divided into two categories. The first involves the initial reading time of the first line. The aim this line analysis is to look for any reading difference between the upward entailing vs. downward entailing contexts, regarding specifically the *numeral* region. If upper bounded ('exact') readings are preferentially computed in upward entailing contexts, and if they are enrichments of basic lower bounded readings, one might expect slower reading times for the numeral in the upward entailing than in the downward entailing context. We note that this is essentially what Breheny et al. (2006) showed in contexts that used a contrastive *all* to bias in favor of the upper-bounded interpretation of *some*.

The second category of predictions involves the initial reading time of the second line after the *ordinal* (*third car*) is read, and the re-reading of earlier material that follows reading of the ordinal. The key assumption 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 positive continuation of the downward entailing sentences, as in (19) (1-B+ 2-C in Table 2):

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

If *two* in (19) is not (yet) upper-bounded in the antecedent of the conditional (20), its truth conditional import may be spelled out as follows:

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

The sentence in (20) 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 (19) is upper bounded (21), the result is coherent, as the following paraphrase makes clear:

```
(21)    In any situation in which John has exactly two cars in the garage, he
parks a third in
        the courtyard.
```

So the continuation in (19) does force an upper bounded reading to be embedded in the antecedent of the conditional. Consider next what happens in the *downward entailing* condition when we have the *neutral continuation.*

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

Clearly, the continuation in (22) does not conflict with the ‘at-least’ interpretation of the numeral, as we may see by considering the following explicit paraphrase:

```
(23)    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 (22) 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 (22), unlike the one in (19), does not lead to a contradiction.

Let us now turn to a downward entailing sentence followed by a *negative* continuation.

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

In spite of differing minimally from (19), 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:

```
(25)    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 automatically be 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 (24) can well be true. Again this mode of reasoning applies to all sentences of this shape. Since (24) allows for a lower bounded interpretation, if our claim is correct, we would not expect an upper bounded interpretation to occur. Thus, (24) should turn out to be easier than (19), in spite of the presence of negation, which is usually considered to make things harder.

Thus, both the negative and the neutral sentences should constitute useful controls. We expect an upper bounded reading to be embedded locally under the scope of a downward entailing semantic function only when forced to do so by the surrounding linguistic or extralinguistic context, and only the positive continuation does so. We expect the comparison between the positive continuation and the negative continuation to be similar to the comparison between the positive continuation and the neutral continuation, while the negative continuation and the neutral continuation are predicted not to differ one from the other. However, according to approaches that accept underspecification or context-dependency as the basis of the meaning of numeral sentences, there is no reason why the polarity of the proposition where the numeral occurs should affect its processing or interpretation.

In sum, as far as reading of the second clause is concerned, our experimental design amounts to two non-independent 2×2 designs. In each we predict an interaction between the *polarity* of the first line and the *type of continuation* of the second line. If the numeral in the first clause is not given an upper-bounded interpretation (in the downward entailing conditions), then the *ordinal* numeral in the positive continuation is the semantic trigger that forces the upper bounded construal of the numeral in the first line. It is conceivable that the reader makes this construal while reading the ordinal numeral of the second clause, resulting in slower reading on the ordinal. Alternatively, or in addition, the reader might be slowed in reading the following (*spillover*) region, or s/he might regress back to the numeral in the first line.

Method

Subjects—The participants were 54 native Italian speakers, between the age of 19 and 29 years old. They were graduate and undergraduate students enrolled in Cognitive Science course at the University of Trento, Rovereto, who took part in the experiment in exchange of course credits. Each participant had normal or corrected-to-normal vision.

Procedure—Participants were told to read the sentences that appeared on a PC 19” CRT screen as soon as they fixated a small white rectangle positioned at the beginning of the first word (the gaze trigger). The main instruction was to read the sentences silently and naturally, neither too fast nor too slowly, in order to answer simple comprehension questions which would follow randomly after reading some sentences. Those questions involved what was said in the first line half the time, and the second line the other half of the time. After the participants read each trial, they had to press a joypad button to move on the next trial or to answer the question. In case no sentence appeared after they fixated the gaze trigger, the experimenter recalibrated the eye-tracking system. The length of the experiment ranged from 30 to 45 minutes per participant, and everyone was offered a small break after the first 15 minutes. At the end of the experiment participants were told about the goals of the investigation. Each participant was run individually.

Eye movements were recorded by an EyeLink-II system (SR Research, Osgoode, ON, Canada) equipped with chin rest. The spatial resolution amounted to 0.01° of visual angle. Recording, made with a sampling rate of 500hz, was binocular, but only data from the left eye has been analyzed. We used an Ijama Monitor of 19 inches with a refresh rate of 85 Hz and a resolution of 1024 × 768. People were placed at 72 cm from it. Saccades and fixations were detected on-line with built-in algorithm by EyeLink and it was used a 9 points calibration. Stimuli were presented with the EyeTrack 0.7.2 software (available at http://www.psych.umass.edu/eyelab/) and the character used was Arial 22 pixels. Computations and statistics were carried by the R package (www.r-project.org).

Materials and design—Six counterbalanced lists, automatically built and randomized by the software, included 24 experimental items, 80 filler sentences and 44 simple comprehension questions (24 on the experimental items and 20 on the fillers). The two experimental factors (*polarity* and *type of continuation*) gave rise to six different conditions (in Table 2, the combinations are 1-A+2-A, 1-A+2-B, 1-A+2-C, 1-B+2-A, 1-B+2-B, and 1-B+2-C), therefore 4 out of 24 experimental phrases appeared in a single condition. Moreover 12 items displayed a conditional clause in the *downward entailing* conditions while the other 12 items contained a universal quantifier restriction in the same conditions, exactly like the first experiment. As shown in Table 3, items can be divided into 16 regions of interest, 10 of which were present in every experimental condition, 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. In each item the first clause was ended by a line break, and the second clause (including the conjunction in the upward entailing conditions and the negation in the negative conditions) appeared on the second line, as shown in Table 3.

Results

Every fixation shorter than 70ms or longer than 700ms was excluded from the analysis (cf. Rayner et al., 1989, for justification of cutoff values in this range). Furthermore, we eliminated trials in which participants skipped three or more words in a single saccade and trials in which the proportion of skipped words was beyond three standard deviations from the mean of skipped words in every trial. We also excluded from the analysis trials presenting lost tracks or fixations that couldn't be considered as normal reading fixations. No more than 3% of trials were left out the analysis. The accuracy of answers to the comprehension questions was high (91%) therefore we may conclude the participants read the sentences carefully and were committed to the task.

We report the results from seven reading measures on fixation duration, two on frequency variables (number of fixations), and four on dichotomous variables (underlying the probability of an event). The first group of indices includes *first fixation duration* (the mean of every first fixation on a word), *gaze 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 either to the right or to the left), *regression-path 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 to the right), *conditioned regression-path time* (equal to the *gaze duration time* plus the time spent re-reading just the preceding word). All the reading measures computed on fixation-duration took into account only regions which received at least one fixation; values of zero were excluded from the analysis. The second group of indices includes *second-pass fixation number* (the count of every second pass fixation made on a word) and *conditioned second-pass fixation number* (the count of every second pass fixation made after the reader encountered the disambiguating word). The third group of indices includes *skipping rate* (the probability that a word was skipped on the first pass), *first-pass regression probability* (the probability that the reader made a first-pass regression from a word), *second-pass fixation probability* (the probability that the reader made at least a second pass fixation on that word) and *conditioned second-pass fixation probability* (the probability that the reader made at least a second pass fixation on that word after reading the disambiguating word; we considered this measure to be more diagnostic of the disambiguating word disrupting reading than a more usual unconditioned measure of second-pass fixation probability). We employed a mixed model to analyze these indices. Mixed-effects models offer several advantages with respect to the traditional statistical techniques of data analysis (cf. Baayen, 2008; Baayen et al., 2008). They are robust with respect to missing data and they have slightly superior power. Second, by estimating the random factors (i.e. subjects and items, in the current study) they provide insight into the full structure of the data. Third, they can be applied to both continuous (linear mixed models) and dichotomous variables such as the occurrence of a regressive eye movement (generalized linear mixed models) (cf. Jaeger, 2008).

Three kinds of Linear Mixed Model were conducted to best analyze these three groups of measures: a Linear Mixed Model based on the Gaussian family distribution for the first group, a Generalized Linear Mixed Model based on the Poisson family distribution for the second group and a Generalized Linear Mixed Model based on the binomial family distribution for the third group (see Jaeger, 2008). These models were conducted by utilizing the *R* platform (www.r-project.org) with the lme4 package developed by Bates et al.

(http://lme4.r-forge.r-project.org). The degrees of freedom of the *t* and *F* tests for the Linear Mixed Model, as calculated in the standard way (number of observations minus fixed parameters), are problematic, for they are an upper-bound and may thereby generate p values too small. We generated confidence intervals by the Markov Chain Monte Carlo sampling (by using the function *pvals.fnc* available from the package, as recommended by Baayen et al., 2008) and we will report the *F* and p values computed over 10000 samples. Furthermore, we excluded from the model all the values lying beyond a 2.5 Standard Deviations cut-off threshold in the standardized residuals distribution, in order to obtain a better fit of the model. The raw means we report, then, are calculated excluding the outliers that generated residuals beyond this threshold.

First line first-pass indices results—Table 4 presents the raw means and the results of the models, for every reading index, conducted over the five relevant regions of the first line. Since first line first-pass indices are influenced only by the polarity factor, the type of continuation factor was left out of the analysis. The most interesting region for our purposes is the numeral region. First, the skipping rate of this region was quite high, so all the first-pass indices are computed on this region only on 60% of the trials. This is not surprising since the numerals we employed are short and high frequency words (e.g. *due*, *tre*, *quattro* which in English are *two*, *three*, *four*), therefore they are often skipped by the reader. However, since the skipping rate did not significantly vary across conditions (DE: 40.8% vs. UE: 37.6%, z = 0.22, p = .8) the first-pass indices on this region were not biased. We expected slower reading times on the numeral region in the upward entailing condition than the downward entailing condition. This difference was observed in Regression-path duration (by 10 ms) and Conditioned regression-path duration (by 14 ms). The latter was fully significant and the former, marginally so. This result suggests that the penalty for the upward entailing condition reflected primarily first-pass re-readings on the word preceding the numeral word, that is the *verb1* region (with a nonsignificant contribution from gaze duration on the numeral itself). Importantly, the slowdown on the numeral in the upward entailing conditions is stable across the type of items, as attested by the similarity of effect sizes and the lack of significant interactions between the polarity and type of item factors (conditioned regression-path duration: downward entailing conditional vs. upward entailing conditional: 247 vs. 258 ms; downward entailing quantifier vs. upward entailing quantifier: 252 vs. 272 ms; F = 1.08, p = 0.3).

Significant differences were also observed in the subject region. Almost all the indices we examined showed significantly higher reading times on this region in the downward entailing condition than in the upward entailing one. The verbal material surrounding this region, however, was not the same in the two types of items. The subject region, in fact, was preceded by *se* (*if*) in the conditional items and followed by *che* (*who*) in the quantified items in the downward entailing condition, hence we must check whether reading time differences are constant across the different types of items. Unlike what we found on the numeral region, the interactions between the two factors were highly significant and the marginal means reverse across the two types of items: *Gaze duration*: DE-conditional vs. UE-conditional: 314 vs. 320 ms, DE-quantified vs. UE-quantified: 304 vs. 247 ms, F = 28.33, p < .001; *conditioned regression-path duration*: DE-conditional vs. UE-quantified: 314 vs. 322 ms, DE-conditional vs. UE -quant: 334 vs. 247 ms, F = 58.78, p < .001. These differences very likely reflect the lexical differences between downward and upward entailing conditions. No other significant main effects were found on other first line regions.

Second line first-pass indices results—Table 5 presents the means and significance tests for first-pass reading time measures on line 2. The core results involve the interaction between polarity and the type of continuation factors. This effect was significant for both the positive vs. neutral continuation and positive vs. negative continuation comparisons, in the regression-path duration computed on the last region (spillover). The interaction for the neutral vs. negative continuation comparison, in contrast, was not significant. No other significant main effects or interactions between the polarity and type of continuation factors were found on the other regions and in the other first-pass indices.

There was an overall advantage for the positive continuation. In the positive vs. neutral continuation comparison, the regression-path duration showed a significant main effect of the type of continuation factor on the object2 and spillover regions, which had higher reading times in the neutral continuation by 253 and 129 ms, as shown in Table 5. It is to be noted that the ordinal region is absent in the neutral continuation and the substantive was always longer to balance the length of the pair "*ordinal* + *object2*". In the positive vs. negative continuation comparison the verb2, ordinal and spillover regions, which contained identical words, displayed higher regression-path duration in the negative continuation by 57, 20, and 174 ms. In the negative vs. neutral continuation comparison, only the verb2 region displayed higher reading times in the negative continuation, by 37 ms, whereas the object2 region was more difficult to read in the neutral continuation, which contained a longer word.

Two different conclusions may be drawn from these results. First, according to the regression-path duration the positive continuation was generally easier to read than the negative and the neutral continuations. Of greater interest, the interactions of type of continuation and polarity provide support for our hypotheses, and motivate looking carefully at the second-pass reading indices.

Second-pass indices results—Table 6 displays the probability and the mean number of second pass fixations on the numeral in the first line. To ensure that such regressions were made after the reader encountered the trigger (i.e. the ordinal in the positive and negative continuations), in the conditioned second pass indices we eliminated all the regressions the reader made before reading the trigger. Consider first the probability of re-reading the numeral in the first line in the positive continuation conditions. Here, in all the second pass indices readers made more regressions towards the numeral in the downward entailing conditions than in the upward entailing ones (see Table 6). If we look at the values of the neutral and negative continuation conditions, in contrast, we see that both conditions pattern alike. 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 upward entailing conditions than in the downward entailing ones. The statistical analysis showed that these differences gave rise to several significant interactions. In the positive vs. neutral continuation comparison we found significant interactions in the second pass fixation probability ($z = 2.62$, $p < .01$), second pass fixation number ($z = 2.52$, $p < .01$) and conditioned second-pass fixation probability ($z = 2.03$, $p < .05$), whereas the conditioned second-pass fixation number was not significant ($z = 1.35$, $p = .2$). In the positive vs. negative continuation comparison we found significant interactions in all the indices: second pass fixation probability ($z = 2.56$, $p < .01$), second-pass fixation number ($z = 2.49$, $p < .01$), conditioned second-pass fixation probability ($z = 3.04$, $p < .01$) and conditioned second pass fixation number ($z = 2.76$, $p < .01$). The fact that most of the conditioned second-pass indices are significant tells us that these effects were caused by the readers' behaviour after they encountered the trigger (ordinal region).

Furthermore, in the positive vs. negative comparison we observed interactions exhibiting the same significant trend in the second-pass reading measures computed on several first- and second-line regions: the subject region (DE-positive vs. UE-positive: 42.7% vs. 33.6%, DE-negative vs. UE-negative: 35.4% vs. 40.3%, $z = 2.36$, $p < .05$), modifier (DE-positive vs. UE-positive: 51.6% vs. 41.6%, DE-negative vs. UE-negative: 45.3% vs. 47.2%, $z = 2.09$, $p < .05$), ordinal region (DE-positive vs. UE-positive: 25.8% vs. 20.1%, DE-negative: vs. UE-negative: 26.4% vs. 32.9%, $z = 2.11$, $p < .05$) and object2 (DE-positive: vs. UE -positive: 21.6% vs.

17.3%, DE-negative vs. UE -negative: 27.8% vs. 37.0%, z = 2.32, p < .05). This confirms the idea that this effect was more pronounced in the positive vs. negative continuation comparison.

Finally, the only significant main effect of the type of item factor was found on the *verb2* region (conditional vs. quantified: 53.5% vs. 64.9%, z = 2.47, p < .01), and the only region displaying a significant interaction between the *polarity* factor and the type of item was the *subject* region (DE-conditional vs. UE-conditional: 40.8% vs. 32.1%, DE-quantified vs. UE-quantified: 38.6% 45.3%, z = 3.27, p < .01). None of the interactions discussed above were influenced by the *type of item*, as attested by the absence of any triple significant interaction on those regions.

Discussion

The first important finding is that the phrases for which the Experiment 1 participants preferred an upper bounded reading for the numeral in an upward entailing context, exhibit an early processing penalty on the numeral region of the first clause in Experiment 2. 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 shows that the polarity of the context is a factor systematically exploited by a reader. Our preferred account of this effect is that, if the embedding context is upward entailing, the upper-bounded meaning (on our approach, the scalar implicature) is computed on line, beginning as soon as possible. If the embedding context is downward entailing, the upper bounded meaning is less likely to be computed, and if it is needed later, a scalar implicature is made when needed. Making the scalar implicature slows reading. Breheny et al. (2006) made essentially the same interpretation of their finding that the region containing *some* was read more slowly in a context that is biased by world knowledge to favor an upper bounded construal, compared to a context that permitted the lower bounded construal.

The requirement for an upper-bounded meaning appeared in the second clause of a sentence. Our positive continuations (e.g., *he parks a third car in the courtyard*; 2-B in Table 2) were designed to require an ‘exact’ interpretation of the numeral in the first clause, and to act as a trigger of reanalysis if that interpretation had not been computed. We expected this reanalysis to result in more processing disruption in the downward entailing condition compared to the upward entailing condition, because the upper-bounded interpretation should be made less often in the former than the latter. We compared processing to two types of control sentences. One, the neutral continuation, lacked the ordinal numeral, which triggers the need to recalculate the meaning of the numeral in the downward entailing condition. The second, the negative continuation, differed by one word from our test sentences (namely, negation). Yet, for semantic reasons, the negative continuation was expected not to force a reinterpretation of the numeral in the downward entailing 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. We thus expected an interaction between the polarity of the first clause and the type of continuation.

Our findings are as follows. First, both the neutral and the negative continuations were overall more difficult to read than the positive one. This might be because of the intrinsic complexity of negation (in fact the *verb2* region, which was more difficult to read in the negative condition, is adjacent to the negative operator *non*) on the one hand, and because of the presence of a novel substantive (e.g., *motorcycle*) in the neutral continuation, on the other. Second, and more importantly, we did find the predicted interaction effects, always significant on the first line numeral region in terms of second-pass probability, second-pass fixation number and conditioned second-pass probability. The same interaction was found in the regression-path duration computed on the last region (spillover), 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 upward and the downward entailing 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 downward entailing context, in the positive continuation, while they behaved in the opposite way in both the neutral and the negative continuations.

Considering the first clause, 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 initial reading the numeral was more difficult to read in the upward entailing condition, while in re-reading the numeral region after reading a positive continuation there were more fixations in the downward entailing condition than in the upward entailing one. In contrast, the readers' behaviour was uniform, across the first and the second-pass, in both the neutral and the negative continuation. In these cases, the numeral always received more first- and second-pass processing in the upward entailing condition. This global picture fits our main claim and its implications for processing, and supports our treatment of upper-bounded construals of numerals as scalar implicatures.

Furthermore, if we look at the positive vs. negative continuation comparison, in which there was only one word to distinguish the two conditions (the negation: *non*), we see that the two-way interaction pattern appears in the subject, modifier, ordinal and object2 region as well.

Finally, the difference in the constructions (type of item factor) significantly influenced only the subject and the verb2 region in some reading measures. 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. This shows that the readers' behaviour with respect to the numeral region was influenced by the semantic structure of the two environments (downward entailing vs. upward entailing), 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.

General Discussion

We set ourselves the following main goals:

```
(26)    to establish that structural factors (entailment properties of the local context) affect
   the interpretation of numerals

(27)    to establish that the lower bounded interpretation of numerals occurs preferentially in
   downward entailing contexts (with respect to minimally different upward entailing ones),
   while the upper bounded one occurs preferentially in upward entailing contexts (with respect
   to minimally different downward entailing ones)

(28)    to determine whether one interpretation of numerals is costly to compute, compared
   to the other.
```

We have addressed these questions by investigating two types of functors, *every* and *if*, both of which are downward entailing in their restriction, and contrasting them with minimally different upward entailing contexts (*and*, in the case of *if*; in the case of *every*, we simply displaced the numeral from the restriction to the scope). The minimal differences between our upward entailing and downward entailing materials makes it implausible that our results may stem from idiosyncratic features of the selected items. Direct evidence favoring (26) and (27) was obtained in an off-line questionnaire, and indirect evidence was obtained through an on-line eye movement experiment which revealed a systematic processing penalty associated with contexts in which one forces readings that go against the generalization in (27). Evidence from eye movements during the reading of a numeral in different contexts provided evidence relevant to (28): readers spent longer regressing from a numeral and re-reading the preceding word when a numeral appeared in an upward entailing than a downward entailing contexts.

As reviewed in the Introduction, evidence already exists that supports generalizations like (26) and (27) as applied to scalar terms other than numerals (*some*, *or*, etc.). The 'strengthened' interpretations of these terms ('some but not all,' 'A or B but not both') correspond to the upper bounded or 'exact' interpretation of numerals and are viewed as scalar implicatures. There is substantial debate about how and why implicatures are made, but many proposals are based on the assumption that scalar implicatures are motivated by the logic of Grice's (1989) maxim of informativeness. One example appears as Chierchia et al's (2008) principle of "Optimize Informativeness" (29):

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

The observation that scalar implicatures are made less often in downward entailing than upward entailing contexts can be traced to the fact that entailment patterns are reversed in downward entailing contexts compared to upward entailing ones, so that an implicature in a downward entailing context would result in a logically weaker statement than retaining the basic meaning.

Our observation that numerals are affected by downward entailing vs. upward entailing contexts in a way very similar to how other scalar terms seem to be affected suggests to us that it is worth considering the possibility that numerals, like other scalar terms, have a basic, lower-bounded, interpretation, and that their 'exact' interpretation results from an implicature, like the strengthened interpretations of other scalar terms. We propose that such consideration is usefully done by referring back to Table 1. It seems to us that our findings are most directly compatible with the asymmetric approach in which the lower-bounded meaning is basic. Such an approach claims that an 'exact' implicature is preferentially made in an upward entailing context, but not in a downward entailing context. Further, on the assumption that making an implicature takes processing resources, the approach finds support in our observation that reading of numerals was slowed in upward entailing contexts, where the implicatures are presumably more likely to be made.

The opposite suggestion (which Breheny, 2008, has made for numerals but not for other scalar terms), claims that the 'exact' meaning is basic and the 'at-least' construal arrives as a subsequent pragmatically driven process that weakens the more basic 'exact' construal. Perhaps this approach can also be integrated by the idea that the 'at-least' construal is favored in a downward entailing context but disfavored in an upward entailing context. This suggestion clearly can account for our observed effects of downward vs. upward entailing contexts on the likelihood of the two interpretations. However, it seems to make exactly the wrong prediction about the processing costs. It predicts that the more basic 'exact' interpretation (in an upward entailing context) is more costly than the subsequent, additional process of pragmatic weakening.

The symmetric approaches listed in Table 1 can be made to account for the observation that 'exact' and 'at-least' interpretations are preferred in different environments. They could do this not by making assumptions about the conditions in which implicatures are added to basic meanings but by making assumptions about how meanings are resolved or specified in different contexts. These assumptions would have to have much in common with the assumptions we have made in accounting for why implicatures are more likely to be made in upward entailing than in downward entailing contexts: the reader or listener tends to arrive at the most informative interpretation of a sentence (cf. the Optimize Informativeness principle of (29)). However, the ambiguity approach, at least, misses what we consider to be an interesting generalization, namely that numerals seem to act like other scalar terms (which are presumably not ambiguous). Further, neither approach has any obvious way of accounting for our observation that reading a numeral in an upward entailing environment seems to be more costly than reading it in a downward entailing environment.

For these reasons, we favor an approach in which the 'at-least' meaning of a numeral is basic and the 'exact' meaning is arrived at as an implicature, with a cost in processing time. We have mentioned one piece of evidence that calls this into question, namely, Huang and Snedeker's (2009) observation of an immediate interpretation of the upper bounded meaning of *two* (as fast as *all*) vs. a delayed interpretation of the upper bounded meaning of *some*. (Recall that their task was to select, and implicitly to look at, the e.g. "girl who had two/three/some/all of the socks.") Huang and Snedeker's preferred account is that the basic meaning of *two* is the upper-bounded reading, while the basic meaning of *some* is the lower-bounded reading, requiring a scalar inference to permit selection of a referent, a conclusion that is clearly at odds with ours.

We do not, at the present time, have a compelling resolution of this conflict. Someone who adopts the Huang and Snedeker account would presumably be obliged to provide an account for why we observed that reading of numerals was disrupted when they occurred in a context that encouraged an upper-bounded interpretation. Similarly, we are obliged to provide an account of the Huang and Snedeker finding. We have no certain account at present; we strongly suspect that additional research using the visual world paradigm will be informative to this debate. In the meantime, we can offer some speculations. One speculation turns on the difference between tasks. Huang and Snedeker's subjects were encouraged to identify the referent of "girl who has two/some of the socks" as quickly as possible. Note that *two* could be taken to refer to the picture with two socks even in the absence of a complement: it would be enough to say *Pick the girl who has two. Some*, on the other hand, must be composed with its complement, *of the socks*, and in its upper-bounded sense, the phrase *the girl with some of the socks* must be evaluated with respect to the picture of the boy who has the rest of the socks. A second speculation, suggested by remarks made by Barner and Bachrach (in press), observes that *two* is frequently used to denote sets with precisely two members, and this precision of reference might facilitate choice of a picture with precisely two objects – especially in the context of a study that repeatedly asks subjects to seek only strengthened meanings (and never to select sets that are consistent with unstrengthened meanings).

These speculations permit us to account for Huang and Snedeker's results while maintaining our position that the meaning of an integer like *two* is the weaker, lower-bounded meaning. It is possible to defend such the position that this is the basic lexical meaning of an integer. For instance, Barner and Bachrach (in press) present arguments that young children's initial, basic, meaning of the numeral *one* is a lower-bounded 'at-least' meaning, the same as the meaning of *a*, and that the child strengthens this to an exact meaning by an implicature based on the contrast with the numeral *two* – an implicature that can be made only when the child has acquired a meaning for *two* as well as *one*. Exact meanings of other numeral terms are similarly created by implicature only when the meanings of larger terms are available.

Another possibility is that the basic lexical meaning of a numeral is in fact the exact value, but that the compositional meaning of some uses of numerals in sentences is the 'at-least' meaning. Consider the fact that noun phrases occur in (at least) two main syntactic structures: after the copula (or in small clauses) and in argument position (e.g., subject, or object, or object of a preposition). In predicate vs. argument position they are interpreted differently: as properties and as entities/quantifiers respectively. Properties and entities/quantifiers are different semantic structures, with different logical types and denotations. These two interpretations are related via mechanisms of compositional semantics. Typically one starts with one of the two (say, the predicative one) and then a function of some sort derives the other in the appropriate spot of the sentence. (cf. Partee, 1987)

With this in mind, here is one way to view numbers: *two cars* starts out as a predicate with an 'exact' interpretation. When it is turned into an argument (as it must be in our materials, but arguably not in the Huang and Snedeker materials), it gets the 'at-least' interpretation via the compositional mechanism that turns predicates into arguments (which is termed 'existential closure' in the semantics literature). Note that the interpretation of (30a), with *two fast cars* in predicate position, is quite clearly the 'exact' reading, while (30b), with the phrase in an argument position, is logically consistent with Simon having driven many fast cars.

(30) a. Those are two fast cars.

b. Simon drove two fast cars.

The 'exact' interpretation for *two cars* in argument position arises, when it does, as an implicature. The point of all this is that there might be an early stage of semantic processing in which you have to go through the 'exact' interpretation, even if by the end of the sentence you wind up with an 'at-least' interpretation. This hypothesis, if correct, might reconcile Huang and Snedeker's findings with ours.

There are many interesting questions remaining about how scalar implicatures are made. One is why implicatures (or more neutrally, 'exact' readings) of numerals are quite frequently made in downward entailing environments, and why basic ('at-least') meanings are quite frequently retained in upward entailing environments. We certainly do not deny that world knowledge, plausibility, discourse context, etc. play a large role in arriving at interpretations, and we acknowledge the need for a processing theory that accounts for how such factors are integrated with the logical-strengthening reasoning that underlies the reasons we have advanced for why implicatures are dispreferred in downward entailing environments. A second, and related, interesting question is whether implicatures are made as a result of global Gricean reasoning, where the informativeness of a complete utterance is maximized, or whether they are made locally, at the level of individual constituents of an utterance (see Chierchia et al., 2008, and Levinson, 2000, for discussion). A closely related question is whether implicatures are made as a result of logical or pragmatic reasoning applied on the fly, or whether this reasoning has been incorporated into the grammar, so that e.g. the grammatical property of being a downward entailing environment directly discourages whatever grammatical operation applies to trigger an implicature (see Chierchia, 2004, and Chierchia et al., 2008, for discussion and for proposal of an explicit grammatical mechanism of 'exhaustification' that brings about what appears to be an implicature). Recall that we proposed our observation that reading is immediately slowed when a numeral is read in an upward entailing context supports the claim that a scalar inference is made at that point. This proposal, if true, clearly argues against a global reasoning account and nicely fits with a fast, grammatically triggered process.

Conclusions

The interpretation of numerals seems to respond to entailment patterns of contexts in ways that are similar to other scalar terms: their upward-bounded, strengthened, interpretations are less preferred in downward entailing than in upward entailing environments. This pattern is observed both in off-line questionnaire data and in on-line eyetracking data. Further, reading of a numeral is slowed in an upward entailing compared to a downward entailing environment, suggesting that there is some processing cost to arriving at the upward-bounded interpretation. We take these observations to be evidence that the basic meaning of a numeral (at least its basic meaning in a sentence context) is the weaker, lower-bounded, 'at-least' construal, and that the stronger, upper-bounded, 'exact' construal is arrived at as a scalar implicature. We further claim that the interpretation of numerals, like the interpretation of other scalar terms, is responsive to principles according to which the informativeness of utterances is optimized, and suggest that these claims encourage the further exploration of the processes by which scalar implicatures are made.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Lyn Frazier, Maria Teresa Guasti, Harald Baayen, Luigi Lombardi, Roberto Cubelli, Francesco Vespignani and Remo Job for their help and insightful suggestions on this work. We are also grateful to the eye-movements lab of the department of cognitive sciences and education, University of Trento, Rovereto (Italy) – especially to Matteo Valsecchi and Elena Betta for their precious help in setting up the experiments, and Angelo Maravita and Francesca Foppolo (University of Milano-Bicocca) for their help in collecting data. Author Clifton's contributions were supported in part by NIH Grant HD-18708 to the University of Massachusetts.

References

Baayen, RH. A Practical Introduction to Statistics Using R. Cambridge, UK: Cambridge University Press; 2008. Analyzing Linguistic Data.

Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language (Special issue on Emerging Data Analysis) 2008;59:390–413.

Barner D, Bachrach A. Inference and exact numerical representation in early language development. Cognitive Psychology. (in press).

Bezuidenhout A, Cutting JC. Literal meaning, minimal propositions, and pragmatic processing. Journal of Pragmatics 2002;34:433–456.

Bott L, Noveck IA. Some utterances are underinformative: The onset and time course of scalar inferences. Journal of Memory and Language 2004;51(3):437–457.

Breheny R, Katsos N, Williams J. Are Scalar Implicatures Generated by Default? Cognition 2006;100 (3):434–463. [PubMed: 16115617]

Breheny R. A new look at the semantic and pragmatics of numerically quantified noun phrases. Journal of Semantics 2008;25:93–140.

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

Chierchia, G.; Crain, S.; Guasti, MT.; Gualmini, A.; Meroni, L. The acquisition of disjunction: evidence for a grammatical view of scalar implicatures. In: Do, AH-J.; Dominguez, L.; Johansen, A., editors.

Proceedings of the 25th annual Boston University conference on language development. Somerville, MA: Cascadilla Press; 2001. p. 157-168.

Chierchia, G.; Spector, B.; Fox, D. The Grammatical View of Scalar Implicatures and the Relationship between Semantics and Pragmatics, to appear. In: Portner, P.; Maienborn, C.; von Heusinger, K., editors. Handbook of Semantics. Mouton de Gruyter; 2008.

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

Crain, S.; Thornton, R. Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics. Cambridge, MA: MIT Press; 1998.

Gazdar, G. Pragmatics: Implicature, presupposition, and logical form. New York: Academic Press; 1979.

Grice, HP. Logic and Conversation. In: Cole, P.; Morgan, JL., editors. Syntax and Semantics, vol. 3: Speech Acts. New York: Academic Press; 1967. p. 41-58.1975. Reprinted in H. P. Grice 1989, (pp. 22–40)

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

Guasti, T.; Gualmini, A.; Meroni, L.; Crain, S.; Foppolo, F.; Chierchia, G. Semantic and Pragmatic Competence in Children's and Adults' Comprehension of *Or*. In: Noveck, I.; Sperber, D., editors. Experimental Pragmatics. New York: Palgrave McMillanHorn, L; 2004. (1972). On the semantic properties of the logical operators in English. Unpublished Doctoral Dissertation, UCLA

Horn, LR. Issues in the investigation of implicature. In: Petrus, K., editor. Volume in honor of Grice. 2007. To appear

Huang Y, Snedeker J. On-line Interpretation of Scalar Quantifiers: Insight into the Semantic-Pragmatics Interface. Cognitive Psychology 2009;58:376–415. [PubMed: 18976987]

Jaeger TF. Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. Journal of Memory and Language. (Special issue on Emerging Data Analysis) 2008;59:434–446.

Ladusaw, B. Unpublished PhD Dissertation, University of Texas at Austin. 1979. Polarity sensitivity as inherent scope relations. (distributed by IULC)

Levinson, S. Presumptive meanings. Cambridge, MA: MIT Press; 2000.

Noveck IA. When children are more logical than adults: Experimental investigations of scalar implicature. Cognition 2001;78(2):165–188. [PubMed: 11074249]

Noveck IA, Chierchia G, Chevaux F, Guelminger R, Sylvestre E. Linguistic-pragmatic factors in interpreting disjunctions. Thinking and Reasoning 2002;8(4):297–326.

Noveck IA, Posada A. Characterizing the time course of an implicature: An evoked potentials study. Brain and Language 2003;85:203–210. [PubMed: 12735938]

Papafragou, A.; Musolino, J. Cognition. Vol. 86. 2003. Scalar implicatures: Experiments at the semantics-pragmatics interface; p. 253-282.

Partee, BH., et al. Noun Phrase Interpretation and Type-Shifting Principles. In: Gorenendijk, J., editor. Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers. Foris: Dordrecht; 1987.

Rayner K. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin 1998;124:372–422. [PubMed: 9849112]

Rayner K, Sereno S, Morris R, Schmauder R, Clifton C Jr. Eye movements and on- line language comprehension processes. Language and Cognitive Processes 1989;4:SI 21–SI 50.

Schwarz, F.; Clifton, C., Jr; Frazier, L. Focus, downward entailing contexts, and scalar implicatures. University of Massachusetts Occasional Papers in Linguistics. (in press)

Sperber, D.; Wilson, D. Relevance: Communication and Cognition. Oxford: Basil Blackwell; 19851995.

Table 1

Approaches to the semantics of numerals

Symmetric approaches	**Asymmetric approaches**
Ambiguity: Numerals are lexically ambiguous between two construals (Horn, 1992)	Lower-bounded basic: The upper-bounded construal is derived (Horn, 1984; Levinson, 2000; Chierchia et al., 2008)
Underspecification: Numerals are under-specified; interpretation is driven by relevance (Sperber & Wilson, 1985/1995; Carston, 198)	Upper-bounded basic: the lower-bounded construal is derived (Breheny, 2008)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Example of a conditional item with the different type of continuations, Experiment 2.

First line	Second line
(1-A) *UE*	(2-A) *neutral continuation*
Giovanni ha **due** macchine in garage e John has **two** cars in the garage and	parcheggia una motocicletta nel cortile esterno. he parks a **motorcycle** in the courtyard.
(1-B) *DE*	(2-B*) positive continuation*
Se Giovanni ha **due** macchine in garage If John has **two** cars in the garage	parcheggia una terza macchina nel cortile esterno. he parks a **third car** in the courtyard.
	(2-C) *negative continuation*
	non parcheggia una terza macchina nel cortile esterno. he **doesn't** park a **third car** in the courtyard.

--note: **boldfaced** terms are the critical terms; they were not shown in boldface to the subjects in the experiment

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 3

First and Second line regions for conditional and quantified items

first line							
Regions	**Start**	**subject**	**Pronoun**	**verb1**	**numeral**	**object1**	**modifier**
	conditional items						
sentence	(se)	Gianni		parcheggia	due	macchine	in garage
Gloss	(if)	John		parks	two	cars	in the garage
	quantified items						
sentence	nel mio quartiere ogni	ragazza	(che)	ha	due	fratelli	più grandi
Gloss	in my neighboorhood every	girl	(who)	has	two	brothers	older
second line							
Regions	And	negation	verb2	determiner	ordinal	object2	spillover
	conditional items						
sentence	(e)	(non)	parcheggia	una	terza/-	macchina/motocicletta	nel cortile esterno
Gloss	(and)	(not)	parks	a	third/-	car/motorcycle	in the courtyard
	quantified items						
sentence	(e)	(non)	desidera	un	terzo/-	fratello/sorella	di età inferiore
Gloss	(and)	(not)	whishes	a	third/-	brother/sister	younger

Table 4

First line first pass indices

Index	Regions				
	Subj	Verb1	Numeral	Object	Modifier
First Fixation Duration					
UE mean	*224*	*279*	*230*	*228*	*265*
DE mean	*231*	*284*	*223*	*233*	*261*
F. val. / p. val.	3.93 / p. < .05	4.96 / p.< .05	3.30 / p < .1	2.95 / p < .1	0.73 / p. > .1
Gaze Duration					
UE mean	*281*	*364*	*248*	*292*	*401*
DE mean	*309*	*362*	*240*	*293*	*391*
F. val. / p. val.	18.83 / p. < .001	2.18 / p. > .1	2.82 / p. < .1	0.12 / p. > .1	1.62 / p. > .1
Regression Path					
UE mean	*296*	*379*	*262*	*320*	*419*
DE mean	*338*	*379*	*252*	*307*	*434*
F. val. / p. val.	30.84 / p. < .001	2.08 / p. > .1	3.34 / p. < .1	2.60 / p. > .1	1.90 / p. > .1
Conditioned Reg. Path					
UE mean	*281*	*365*	*264*	*313*	*418*
DE mean	*323*	*368*	*250*	*306*	*412*
F. val. / p. val.	39.80 / p. < .001	4.68 / p. < .05	6.85 / p. < .01	0.76 / p. > .1	0.29 / p. > .1

Table 5

Regression-path duration computed on second line regions

Continuation means	**Regions**			
	verb2	**ordinal**	**object2**	**spillover**
Positive				
UE mean	*432*	*259*	*245*	*1200*
DE mean	*434*	*248*	*233*	*1362*
Neutral				
UE mean	*448*	-	*492*	*1466*
DE mean	*456*	-	*493*	*1355*
Negative				
UE mean	*494*	*267*	*249*	*1560*
DE mean	*487*	*280*	*259*	*1350*
Type of Continuation	verb2	ordinal	object2	spillover
Positive vs. Neutral				
F. val. / p. val.	1.63 / p. > .1		727 / p. < .001	6.59 / p. < .05
Positive vs. Negative				
F. val. / p. val.	45.96 / p. < .001	18.61 / p. < . 001	2.27 / p. > .1	7.54 / p. < .01
Neutral vs. Negative				
F. val. / p. val.	26.72 / p. < .001		506 / p. < .001	0.32 / p. > .1
continuation*polarity	verb2	ordinal	object2	spillover
Positive vs. Neutral				
F. val. / p. val.	1.93 / p. > .1		0.19 / p. > .1	7.48 / p. < .01
Positive vs. Negative				
F. val. / p. val.	.01 / p. > .1	0.68 / p. > .1	0.82 / p. > .1	10.99 / p. <.001
Neutral vs. Negative				
F. val. / p. val.	0.64 / p. > .1		0.56 7 p. > .1	0.72 / p. > .1

Table 6

Second pass indices computed on numeral region

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