When *Three* Is Not *Some*: On the Pragmatics of Numerals

Einat Shetreet[1,2], **Gennaro Chierchia**[2], **and Nadine Gaab**[1,3,4]

Abstract

■ Both numerals and quantifiers (like *some*) have more than one possible interpretation (i.e., weak and strong interpretations). Some studies have found similar behavior for numerals and quantifiers, whereas others have shown critical differences. It is, therefore, debated whether they are processed in the same way. A previous fMRI investigation showed that the left inferior frontal gyrus is linked to the computation of the strong interpretation of quantifiers (derived by a scalar implicature) and that the left middle frontal gyrus and the medial frontal gyrus are linked to processing the mismatch between the strong interpretation of quantifiers and the context in which they are presented. In the current study, we attempted to characterize the similarities and differences between numbers and quantifiers by examining brain activation patterns related to the processing of numerals in these brain regions. When numbers were presented in a mismatch context (i.e., where their strong interpretation did not match the context), they elicited brain activations similar to those previously observed with quantifiers in the same context type. Conversely, in a match context (i.e., where both interpretations of the scalar item matched the context), numbers elicited a different activation pattern than the one observed with quantifiers: Left inferior frontal gyrus activations in response to the match condition showed decrease for numbers (but not for quantifiers). Our results support previous findings suggesting that, although they share some features, numbers and quantifiers are processed differently. We discuss our results in light of various theoretical approaches linked to the representation of numerals. ■

INTRODUCTION

The comprehension of scalar expressions has been extensively studied in logic, linguistics, and psycholinguistics (e.g., Shetreet, Chierchia, & Gaab, in press; Huang & Snedeker, 2009a, 2009b; Chierchia, Fox, & Spector, 2008; Chierchia, 2004; Noveck, 2001; Grice, 1975; Horn, 1972). These studies usually focus on the interpretation of weak scalar expressions. Scalar expressions form ordered scales with other scalar expressions of the same type[1] (e.g., ⟨some, many, most, every/all⟩; ⟨or, and⟩; ⟨may, must⟩; and ⟨warm, hot, boiling⟩). Weak scalar expressions, such as *some*, which is positioned at the lower end of the scale, have two possible interpretations, illustrated in Examples (1) and (2). Under the lexical (logical) meaning, the weak scalar expression is compatible with the strong member of the same scale (e.g., *some* is compatible with *every/all*, or in other words, its logical meaning is *at least some*, as shown in Example (2)). However, the interpretation of the weak scalar is often strengthened to exclude the strong member (e.g., *some but not all*, as can be seen in Example (1)).

(1) Some dogs in this neighborhood bark at night.
(2) If some dogs in this neighborhood bark at night, we won't get enough sleep.
(3) Every dog in this neighborhood barks at night.

Although it is controversial what linguistic component drives the interpretation strengthening of weak scalar expressions (e.g., grammatical or pragmatic, for more details, see Chierchia, 2004; Noveck & Sperber, 2007), there is a general agreement regarding the computation process that generates the strong interpretation. This computation occurs through a scalar implicature by considering the alternatives of the weak scalar expression like the quantifier *some*. When encountering this quantifier, language users consider the other members of the quantifier scale as alternatives for what they have encountered (e.g., Example (3) as an alternative for Example (1)). Using the weak scalar and assuming that the speaker is cooperative and knowledgeable indicates to the listener that the strong alternative (i.e., Example (3)) does not hold. Thus, the listener interprets *some* using the strong (pragmatic) interpretation of "*some but not all*."

Like quantifiers, numerals are also ordered on a scale (⟨one, two, three…⟩). They also have more than one interpretation as shown in Examples (4) and (5). In Example (4), the numeral *three* has the strong interpretation of *three and not more* (or in other words *exactly three*), which corresponds to the strong interpretation of quantifiers (i.e., *some and not all*). In Example (5), *three* has the weak interpretation of *at least three*, which corresponds to the weak interpretation of quantifiers (i.e., *some and possibly all*). It has therefore been argued by the classical neo-Gricean approach that the same computation process applies to quantifiers and numerals (e.g., van Rooij & Schulz, 2006; Horn, 1972): A scalar implicature is computed with a weak scalar by considering its

---

[1]Boston Children's Hospital, [2]Harvard University, [3]Harvard Medical School, [4]Harvard Graduate School of Education

Journal of Cognitive Neuroscience X:Y, pp. 1–10
doi:10.1162/jocn_a_00514

alternatives, followed by the inference that the strong alternative does not hold, and thus gives rise to the strong interpretation (i.e., the *some but not all* interpretation in (1) and the *exactly three* interpretation in (4)).

(4) Three dogs in this neighborhood bark at night.
(5) If three dogs in this neighborhood bark at night, we won't get enough sleep.

Interestingly, the interpretation of both quantifiers and numbers is sensitive to the grammatical aspects of the sentences in which they appear. Examples (1) and (4) present upward entailing contexts[2] in which the interpretation of weak scalar expressions is usually strengthened. In Examples (2) and (5), the scalar expressions appear in downward entailing contexts in which the weak interpretation is preferred. Experimental results confirmed this grammatical effect on weak quantifiers and weak numbers (Panizza, Chierchia, & Clifton, 2009; Panizza, Huang, Chierchia, & Snedeker, 2009). Such results support the idea that the two scalar expressions are processed by the same mechanism.

However, despite these similarities, differences between quantifiers and numerals are present in their linguistic behavior as well as in experimental findings. Unlike quantifiers, which have the "at least" or "exact" interpretations (i.e., *some and possibly all* and *some but not all*, respectively), numerals also have an "at most" interpretation (e.g., Breheny, 2008; Horn, 2004; Carston, 1998; although see Panizza & Chierchia, 2011, for an alternative explanation for this interpretation). Example (6), borrowed from Musolino (2004), illustrates the "at most" interpretation. Here, the reader is likely to interpret *three mistakes* as *at most three mistakes*. Using similar statement in a Truth Value Judgment Task, Musolino (2004) showed that adult participants get the "at most" interpretation of numerals (as well as the "at least" interpretation). This was further replicated in other studies (e.g., Geurts, Katsos, Cummins, Moons, & Noordman, 2010).

(6) You can make three mistakes and still pass the test.

Quantifiers and numbers show a different behavior also in experimental investigations of language processing and language acquisition. In eye tracking experiments, look shifts to targets for sentences containing *some* were delayed compared with sentences containing *all*, but not for sentences containing *two* (which referred to the subset, like *some*) compared with sentences containing *three* (which referred to the total set, like *all*; Huang & Snedeker, 2009a, 2009b). In a dual-task paradigm, high memory load resulted in fewer strong interpretation responses for quantifiers but more strong interpretation responses for numerals (Marty, Chemla, & Spector, 2013).

Children also exhibit different behavior with quantifiers and numbers. Several studies using various tasks have shown that children arrive at the strong ("exact") interpretation of numerals more often and more rapidly than they arrive at the strong interpretation of quantifiers (Huang & Snedeker, 2009b; Hurewitz, Papafragou, Gleitman, & Gelman, 2006; Papafragou & Musolino, 2003; Noveck, 2001). For example, Papafragou and Musolino (2003) used sentences such as "some/two horses jumped over the fence" with a scenario in which all three horses jumped over the fence and asked their participants to judge the truth value of the sentence. In such mismatch cases, both true and false responses are acceptable, as the truth value depends on the interpretation of the scalar expression. Five-year-old children gave more adult-like responses (by rejecting the sentences as false) for sentences containing weak numbers than for sentences containing weak quantifiers (also see Noveck, 2001).

These linguistic and experimental differences between quantifiers and numerals have led researchers to suggest that quantifiers and numerals are represented and processed differently (e.g., Breheny, 2008; Carston, 1998). For example, it has been suggested that the lexical meaning of numbers is the "exact" interpretation, which corresponds to the strong interpretation for the quantifiers (Breheny, 2008). Under this account, the "at least" interpretation (which corresponds to the weak underived interpretation of the quantifiers) is derived through a pragmatic process different from the scalar implicature, which is assumed with quantifiers.

Our study attempts to characterize the similarities and differences between numerals and quantifiers using neuroimaging techniques. Our investigation is based on a previous fMRI study that examined the neural correlates of processing scalar implicatures with the weak quantifier *some* (Shetreet et al., in press). That study focused on similarities and differences between mismatched scalar implicature (e.g., the sentence "some giraffes have balloons" presented with a picture in which all of the giraffes had balloons), matched scalar implicature (e.g., same sentence presented with a picture in which only some of the giraffes had balloons), and no implicature (e.g., using sentences including *every* with both picture types) conditions (see Table 1). That study aimed to dissociate the processing of scalar implicature computation from the processing of mismatch between scalar implicatures and the context in which they were presented. Both mismatched and matched implicatures with the quantifier *some* showed increased activation in the left inferior frontal gyrus (IFG; Brodmann's area [BA] 47) when compared with no implicatures (Table 1). Thus, this region, which has a well-known role in semantic processing (Hagoort, Baggio, & Willems, 2009; Hagoort, 2005; Sakai, 2005; Homae, Hashimoto, Nakajima, Miyashita, & Sakai, 2002; Dapretto & Bookheimer, 1999; see also meta-analysis studies, e.g., Binder, Desai, Graves, & Conant, 2009; Bookheimer, 2002; Fiez, 1997), was linked to implicature computation. By contrast, mismatched, but not matched, implicatures showed increased activations in the left anterior middle frontal gyrus (MFG)

Table 1. Summary of Conditions and Results from Shetreet et al. (in press)

Conditions		
Condition	*Sentence*	*Picture*
Mismatch weak quantifier (*some*ALL)	“some giraffes have balloons”	ALL the giraffes have balloons
Match weak quantifier (*some*SOME)	“some mice have grapes”	SOME mice have grapes
Strong quantifier (*every*ALL)	“every rabbit has keys”	ALL the rabbits have keys
(*every*SOME)	“every cat has flowers”	SOME cats have flowers

Main Results	
Comparison	*Activated Region*
Implicature Generation: someALL > *every*ALL and *some*SOME > *every*ALL	Left IFG (BA 47)
Implicature mismatch: *some*ALL > *some*SOME[a]	Left MFG (BA 10)
	Left MeFG

[a] In the analysis reported in the article, a *some*NONE condition was included in this conjunction as well, to control for picture and response types.

and the medial frontal gyrus (MeFG)/ACC (Table 1). Thus, these prefrontal regions were linked to the processing of the mismatch between the implicature and the context (i.e., a picture). Interestingly, these regions have been linked to high cognitive functions, such as conflict monitoring, cognitive control, and truth value judgment (Wolfensteller & von Cramon, 2011; Mansouri, Tanaka, & Buckley, 2009; Wendelken, Nakhabenko, Donohue, Carter, & Bunge, 2008; Carter & van Veen, 2007).

The main goal of the current study was to test the hypotheses of the neo-Gricean approach to numerals by examining brain activations in response to numerals in the same brain regions that were previously linked to the processing of scalar implicatures with quantifiers, left IFG, left MFG, and MeFG/ACC. We sought to determine whether, as assumed by the classical neo-Gricean approach, the representation and the derivation of the interpretations of numerals and quantifiers are performed by the exact same mechanism (van Rooij & Schulz, 2006; Horn, 1972) or not. This approach predicts that brain activations in response to sentences with weak numerals should show substantial similarities to the brain activations observed with sentences with weak quantifiers in Shetreet et al. (in press). Specifically, increased activations are predicted for weak numerals in the left IFG, regardless of the context in which they appear. Additionally, increased activations in the left MFG and MeFG/ACC are predicted for weak numerals presented in a mismatch context, but not for weak numerals presented in a match context. As mentioned above, several other approaches have been suggested for the representation of numerals. In our final discussion, we consider the predictions of the different approaches, although our conclusions do not allow us to settle matters of contention among them. Generally, the alternative approaches assume that numerals and quantifiers are processed differently and therefore predict that numerals would not elicit activation pattern similar to the one observed with quantifiers. If the two scalar expressions are represented and processed differently, they should mainly differ on implicature generation, which was linked to the left IFG (Shetreet et al., in press).

METHODS

Participants

Thirteen adult right-handed native English speakers without neurological, hearing, or language impairment participated in the study (eight women, ages = 19–30 years, mean = 23.4 years). These were the same participants that took part in the experiment reported in Shetreet et al. (in press). All the participants gave informed consent before the experiment and were compensated for their participation. Two additional participants were excluded from the experiment: one due to technical problems (the monitor in the MRI machine was not adjusted properly so the participant did not have a clear view of the pictures) and the other due to awareness to the experimental goal (the participant had extensive knowledge of linguistic theory and therefore was aware of the two possible interpretations of scalar terms and the process by which they are derived). The study was approved by the Institutional Review Board of Boston Children’s Hospital.

Materials and Procedure

This fMRI experiment comprised a simple meaning-matching task, as described in Shetreet et al. (in press), and a control voice-matching task. In the experimental meaning-matching task, participants heard a sentence and saw a picture and were asked to decide if the sentence matches the picture. Sentences included a number (*three*

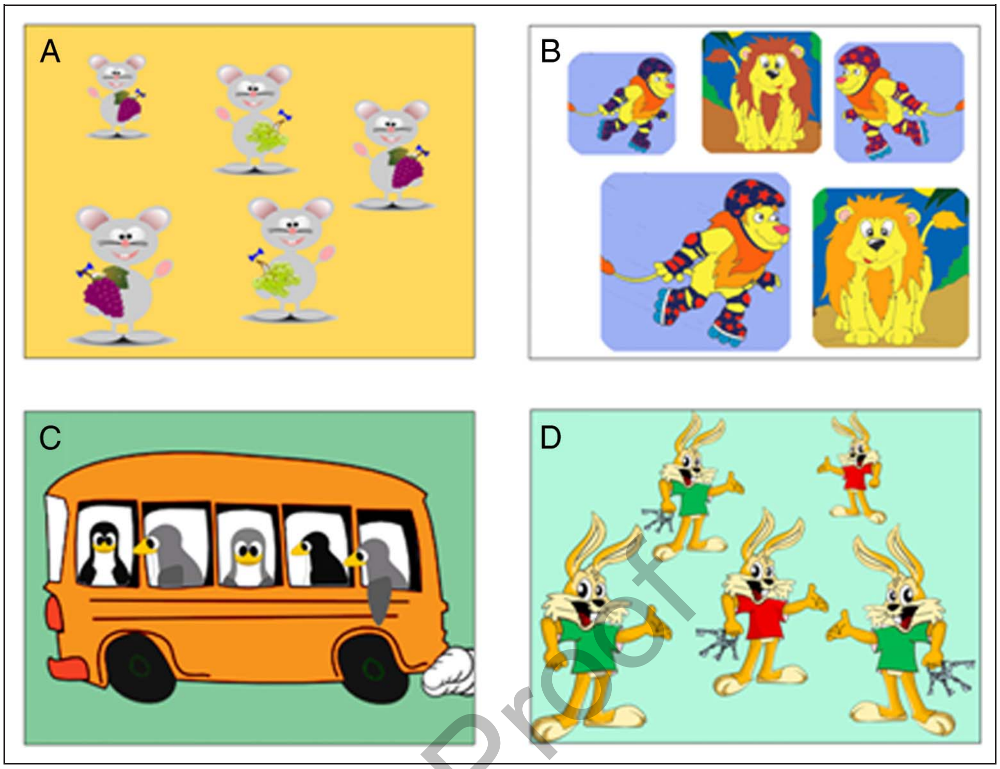

Figure 1. Examples for pictures used in the experimental task (meaning-matching task). Picture (A) was presented with the sentence "three mice have grapes" (for the weak number mismatch condition), Picture (B) with "three lions are skating" (for the weak number match condition), Picture (C) with "five penguins are on the bus," and Picture (D) with "five rabbits have keys" (for the strong number conditions).

or *five*) with a noun at the subject position. All the words in the sentences had age of acquisition before 5 years (as determined by the McArthur-Bates communicative development inventories; Dale & Fenson, 1996). The words were balanced across conditions by using the same words in different combinations (e.g., "three elephants are drinking" with "five elephants are dancing" or "five giraffes are drinking"). Sentences were presented auditorily. A female native American English speaker recorded the sentences in random across conditions. There were no differences in the duration of the sentences among the conditions, $F(3, 69) = .46, p = .71$. The sentences described an action performed by the subjects (e.g., "three elephants are drinking"), the location of the subjects (e.g., "three zebras are on the boat"), or a possession of the subjects (e.g., "three giraffes have balloons").

The pictures to be matched to the sentences included five individuals of the same type (e.g., five giraffes or five girls). In half of the pictures, all five individuals had the same property that was stated in the sentence (e.g., five mice with grapes; Figure 1A), and in the other half, three of the individuals had the same property (e.g., three skating lions; Figure 1B). The combination of sentence type (*three* or *five*) with the picture type (THREE or FIVE) produced four conditions, following the critical quantifier conditions from Shetreet et al. (Table 2 and Figure 1): (1) the weak number mismatch condition with *three* sentences and FIVE pictures (*three*FIVE), corresponding to the implicature mismatch condition with the quantifiers (the *some*ALL condition with *some* sentences and ALL pictures); (2) the weak number match condition with *three* sentences and THREE pictures (*three*THREE), corresponding to the matched implicature condition with the quantifiers (the *some*SOME condition); (3) the strong number true condition with *five* sentences and FIVE pictures (*five*FIVE); and (4) the strong number false condition with *five* sentences and THREE pictures (*five*THREE). The five sentence conditions (3 and 4) corresponded to

Table 2. Number Conditions Used in the Experiment and Their Quantifiers Equivalents (from Shetreet et al., in press)

Condition	*Sentence*	*Picture*	*Expected Response*	*Equivalent Quantifier Condition*[a]
Weak number mismatch: *threeFIVE*	three	FIVE	match/no match	Weak quantifier mismatch: *some*ALL
Weak number match: *three*THREE	three	THREE	match	Weak quantifier match: *some*SOME
Strong number true: *five*FIVE	five	FIVE	match	Strong quantifier true: *every*ALL
Strong number false: *five*THREE	five	THREE	no match	Strong quantifier false: *every*SOME

[a] Based on Shetreet et al. (in press).

the no implicature condition with the quantifier "*every*" (*every*ALL and *every*SOME, respectively). Each condition was sampled 20 times (resulting in a total of 80 sentences). In each run, these conditions were intermixed with the quantifier conditions, which are thoroughly discussed elsewhere (Shetreet et al., in press).

We also included a control task to serve as a baseline for auditory, lexical, and syntactic activations, as well as for activations related to decision-making (for the ROI analysis).[3] The control task included a voice-matching task that required only a shallow processing of the sentence, as the participants had to focus on the voice of the speaker, rather than on the content of the sentence. For this task, we chose a random sample of the sentences that were used in the experimental meaning-matching task. We asked the participants to match the voice of the speaker/s of the sentence to a picture of the speaker/s. The sentences were spoken by a woman, an "alien," or both. The alien voice was created by an audio manipulation in GoldWave program. The pictures in this task included a detailed scene with a woman, an alien, or both. For example, when presented with a picture of a woman and an alien (Figure 2), participants were expected to respond with "a match" if the sentence spoken by a woman were the last word spoken by an alien, but with "a no-match" when the sentence was spoken only by a woman. All the sentences including the alien voice were fillers and were not included in the analysis. Each picture and each sentence were presented once for each participant. There were 20 control sentences and 20 fillers.

In both the experimental and control tasks, each picture was presented for 4 sec. The sentence was played with the initial display of the picture, and a response was required after the ending of the sentence. Rest trials with fixation cross were also included and displayed for 4 sec.

Figure 2. An example for a picture used in the control task (voice-matching task). This picture shows a woman and an alien and was presented with a sentence "five penguins are on the **bus**" where the last word (*bus*) was spoken by "an alien" (a voice manipulation) and the rest were spoken by a woman (for a match response).

An event-related design was used, and the events were presented in random order within each task as determined by optseq software (www.freesurfer.net/optseq). The experimental task was presented in two separate runs (approximately 7.5 min) with 10 trials for each condition in each run. The control task was presented in a single run (approximately 4.5 min). The presentation order of the runs was counterbalanced between subjects. Stimuli were delivered to the participants using Presentation software (nbs.neuro-bs.com). All responses and RTs were recorded.

Before the MRI scan, participants watched an instruction video for each task and practiced the tasks with sentences and pictures that were not used in the MRI experiment. Mismatch scenarios were not included in the practice. Participants did not show any difficulties with the tasks, and feedback about their performance was given to them after the practice. Participants performed other language tasks inside and outside of the MRI, including the intermixed trials of the quantifier sentences. The entire MRI session, including anatomical scans, lasted approximately 1 hr.

Data Acquisition

MRI scans were conducted in a whole-body 3 T, SIEMENS 3T Trio MR scanner. fMRI was performed using a gradient-echo T2*-weighted EPI interleaved sequence with 227 whole-brain images in each run. Thirty-two sagittal slices, 4 mm thick, covering the whole of the cerebrum and most of the cerebellum, were selected. Our acquisition parameters were as follows: field of view = 192, matrix size = 64 × 64, repetition time = 2000 msec, echo time = 30 msec, and flip angle = 90°.

Data Analysis

Image analysis was performed using SPM8 (Wellcome Department of Cognitive Neurology, www.fil.ion.ucl.ac.uk/spm/). Functional images from each participant were slice-time corrected for interleaved acquisition, motion-corrected, normalized to the SPM EPI template, and spatially smoothed using a Gaussian filter (4-mm kernel). Each participant's data were analyzed using a general linear model (Friston et al., 1995) and high-pass filtered at 128 sec. Events were modeled with the onset of the sentence/picture (which was the same) and with the duration of the entire trial. To control for different responses for the *three*FIVE condition (see Results), as well as to follow the analysis from Shetreet et al. (in press), we added a covariate with the responses for this condition. Head motion parameters were added as regressors (Friston et al., 1995). For the group level analyses, one-sample t tests were computed using the individual contrast images. Analyses were carried out with the threshold of $p < .005$, cluster size of $k > 50$ voxels, and cluster level correction of $p < .05$.

Table 3. Areas of Activations in the Different Comparisons Performed ($p < .005$, Cluster Size of $k > 50$ Voxels, and Cluster Level FWE Correction of $p < .05$)

Region	*x*	*y*	*z*	*K*	*t Max*
Weak Number Mismatch vs. Strong Number (threeFIVE > fiveFIVE)					
Left inferior parietal lobule	−48	−49	54	64	6.11
Weak Number Mismatch vs. Weak Number Match (threeFIVE > threeTHREE)					
Left IFG (BA 47)	−48	17	6	51	5.18

Other comparisons did not yield any activation in this threshold.

We also performed an ROI analysis with the regions of activations observed in Shetreet et al. (in press). Three ROIs were selected: (1) The "implicature generation conjunction" (a conjunction of the quantifier mismatched (*some*ALL) and matched conditions (*some*SOME) with a no implicature condition (*every*ALL)) was specially designed to identify regions that participate in implicature generation and not in implicature mismatch processing. Thus, the left IFG (BA 47) ROI in the current study was defined based on this conjunction (Figure 4). (2) The "implicature mismatch conjunction" (a conjunction of the quantifier mismatch condition [*some*ALL] with the other two *some* conditions included in the experiment [*some*SOME and *some*NONE]) was specially designed to identify regions that participate in the processing of the mismatch between the implicature and the context. Thus, the left anterior MFG and the MeFG/ACC ROIs were defined based on this conjunction (Figure 4). For this analysis, we included in our design the control task and computed contrast images of each condition versus the control condition. Average contrast estimates for the four contrasts were extracted from the ROIs using MarsBar (marsbar.sourceforge.net). Planned comparisons, based on the comparisons performed in Shetreet et al. (in press), were computed (see Table 4).

RESULTS

In-scanner Performance

Accuracy rates on the weak number match condition (*three*THREE), the strong number true condition (*five*FIVE), and the strong number false condition (*five*THREE) were above 80% (mean accuracy = 92.6%, *SD* = 4.9). For the weak number mismatch condition (*three*FIVE), 11 participants gave mostly pragmatic responses, choosing the no-match response, and two participants gave logical responses, choosing the match response. When compared with the results from Shetreet et al. (in press), we found more logical responses for the quantifier mismatch condition (with five logical responders) than for the number mismatch condition (with two logical responders). A t test showed that this difference was significant, $t(12) = 2.71$, $p = .02$. Importantly, the same participants gave logical responses for the quantifier condition, but pragmatic responses for the number condition. We also tested the difference in RTs in the four number conditions and found no significant effect, $F(3, 33) = 1.6, p = .16$.

fMRI Results

Whole Brain Analysis

The whole brain analysis for numerals showed a different activation pattern than the one reported for quantifiers in Shetreet et al. (in press). Comparing the weak numeral mismatch *three*FIVE condition with the strong numeral *five*FIVE condition (*three*FIVE > *five*FIVE) revealed activation in the left inferior parietal lobule (BA 40; Table 3). Furthermore, the comparison of the weak numeral mismatch condition with the strong numeral *five*THREE condition (*three*FIVE > *five*THREE) revealed no activation when applying cluster level correction.

The comparison between the weak numeral mismatch *three*FIVE and the weak numeral match *three*THREE conditions (*three*FIVE > *three*THREE) showed increased activation in the left IFG (BA 47; Figure 3 and Table 3), which was linked to scalar implicature generation based on the quantifier conditions (Shetreet et al., in press). Comparing the weak numeral match condition to each of the strong numeral conditions showed no significant activation differences.

ROI Analysis

We also performed an ROI analysis, examining the activations related to the processing of numbers in the regions

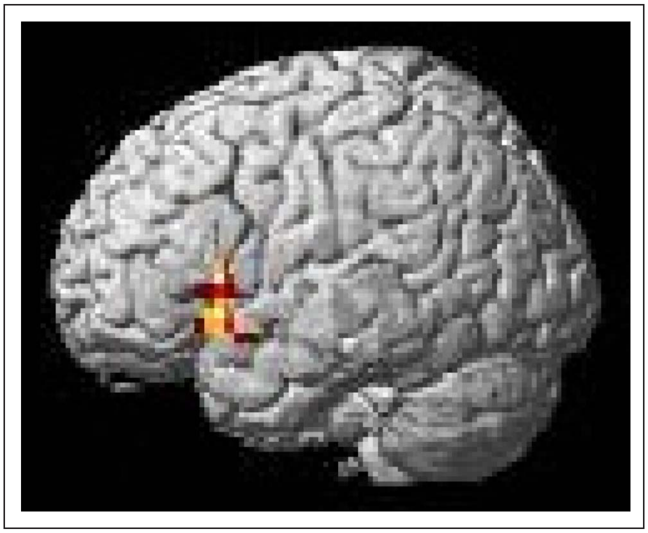

Figure 3. Activation in the left IFG (BA 47), when comparing the weak number mismatch condition with the weak number match condition (*three*FIVE > *three*THREE).

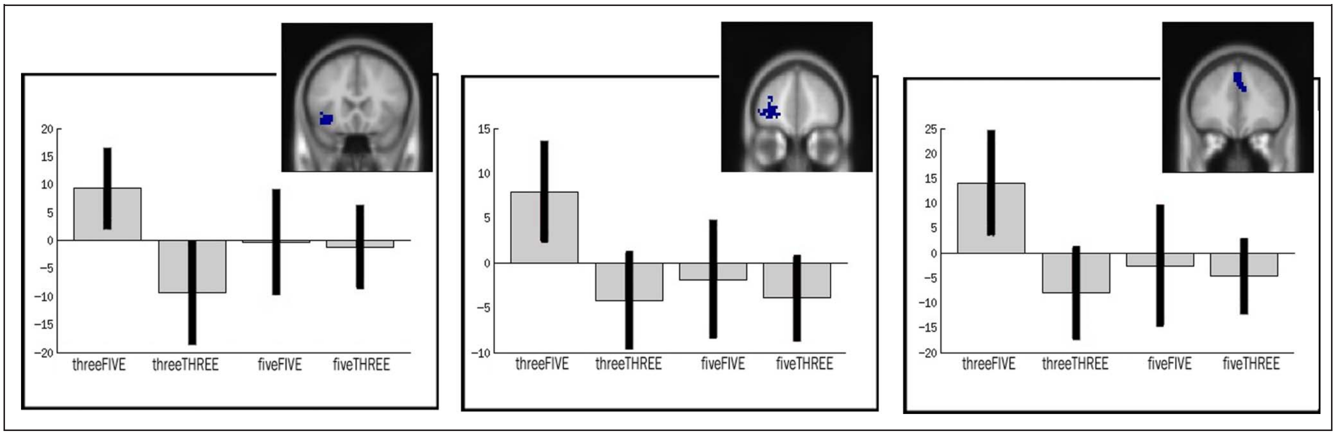

Figure 4. Contrast estimates extracted from the left IFG (left chart), left anterior MFG (middle chart), and MeFG/ACC (right chart). Note that the contrast estimates were defined based on a contrast between each condition and the control condition. That is, all conditions are compared with the same baseline. Therefore, the difference between the conditions should be considered, rather than their absolute values.

observed with the processing of scala implicatures with quantifiers. This ROI analysis allowed us to focus on specific regions defined by a priori hypotheses, thus eliminating the need to correct for a variety of multiple comparisons. First, we examined the activations in the left IFG (BA 47). This region showed increased activation for the weak quantifier mismatched and matched conditions (when compared with the strong quantifier [no implicature] conditions, *some*ALL > *every*ALL and *some*SOME > *every*ALL; Shetreet et al., in press; Tables 1 and 4). Thus, if numerals and quantifiers are processed by the same mechanism, as predicted by the classical neo-Gricean approach, the comparison of weak number mismatch condition (*three*FIVE) with the strong number conditions (*five*FIVE and *five*THREE) and of the weak number matched condition (*three*THREE) with the strong number conditions should yield similar results to the results observed for the quantifiers. Repeated-measures ANOVA of all the number conditions estimates in the left IFG (BA 47) showed a significant main effect, $F(3, 36) = 5.66$, $p = .002$. Using planned comparisons, we observed increased activation for the weak number mismatch condition compared with the strong number conditions, $F(1, 12) = 6.89$, $p = .02$ (Figure 4). However, the weak number match condition showed *decreased* activation when compared with the strong number conditions, $F(1, 12) = 5.31$, $p = .04$. Additionally, similar to the whole brain analysis, the weak number mismatch condition showed greater activation in the left IFG when compared with the weak number match condition, $F(1, 12) = 17.8$, $p = .001$. This pattern of results shows some similarities between the weak number and weak quantifier mismatch conditions, but differences between the weak number and the weak quantifier match conditions (Table 4).

The left MFG and the MeFG/ACC showed increased activation for the weak quantifier mismatch condition when compared with the weak quantifier match condition (*some*ALL > *some*SOME) and with the strong quantifier conditions (*some*ALL > *every*ALL and *every*SOME; Table 4). Additionally, the weak quantifier match condition did not differ from the strong quantifier (no implicature) conditions in these regions (*some*SOME vs. *every*ALL and *every*SOME). Thus, we were specifically interested in comparing the weak number mismatch condition with the weak number match condition and with the strong number conditions, as well as the weak number match condition with the strong number conditions. Both the neo-Gricean approach and its alternatives predict similarities between numerals and quantifiers in this analysis.

Table 4. A Comparison of Quantifiers (Based on Shetreet et al., in press) and Numerals (as Observed in the ROI Analysis in the Current Study)

Contrast	*Quantifiers (Adaptation from Shetreet et al., in press)*	*Numbers (Based on the ROI Analysis)*
Weak mismatch > Strong	Increased activation in left IFG (BA 47)	Increased activation in left IFG (BA 47)
	Increased activation in left MFG and MeFG/ACC	Increased activation in left MFG and MeFG/ACC
Weak match > Strong	Increased activation in left IFG (BA 47)	*Decreased activation in left IFG (BA 47)*
Weak mismatch > Weak match	Increased activation in left MFG and MeFG/ACC	Increased activation in left MFG and MeFG/ACC

Repeated-measures ANOVA showed a significant main effect in both regions [$F(3, 36) = 3.33$, $p = .03$ and $F(3, 36) = 4.71$, $p = .007$ for the left MFG and the MeFG/ACC, respectively; Figure 4]. In both the left MFG and the MeFG/ACC, the weak number mismatch condition showed increased activation when compared with the weak number match condition [$F(1, 12) = 23.05$, $p < .001$ and $F(1, 12) = 16.99$, $p = .001$, respectively], as well as when compared with the strong number conditions [$F(1, 12) = 9.3$, $p = .01$ and $F(1, 12) = 9.52$, $p = .009$, respectively]. Additionally, the weak number match did not show a significant difference from the strong number conditions [$F(1, 12) = 0.001$, $p = .9$ and $F(1, 12) = 0.15$, $p = .7$, respectively]. Unlike the left IFG, which showed differences between numbers and quantifiers, the two scalar expressions showed similar pattern in the left MFG and the MeFG/ACC (also see Table 4).

DISCUSSION

Our study used neuroimaging to explore the processing of sentences with weak numerals in brain regions that were previously observed when processing sentences with weak quantifiers. Our results (as can be seen in Table 4) confirm findings emerging from linguistics and from behavioral studies showing similarities between the two types of scalar expressions in some respects (Panizza, Chierchia, et al., 2009; Panizza, Huang, et al., 2009) but differences in other respects (Huang & Snedeker, 2009a, 2009b; Hurewitz et al., 2006; Musolino, 2004). Our study specifically focused on brain regions that were activated in response to weak quantifiers when presented with mismatched and matched pictures (the left IFG, the left MFG, and MeFG/ACC; Shetreet et al., in press) and considered this pattern of activation as the starting point for the analysis related to weak numerals. We showed that the pattern of brain activations related to weak numerals and weak quantifiers was similar when presented in a mismatch context (i.e., *some*ALL in Shetreet et al., in press, and *three*FIVE in the current study). The weak quantifier mismatch condition led to increased activation in the left IFG (BA 47) and the left MFG and MeFG/ACC when compared with the strong quantifier (no implicature) conditions (Shetreet et al., in press). Likewise, the weak number mismatch condition revealed increased activations in these three brain regions when compared with the strong number conditions. By contrast, weak numerals and weak quantifiers differed on their pattern of brain activations when presented in a matched context (i.e., *some*SOME in Shetreet et al., in press, and *three*THREE in the current study). Whereas the weak quantifier match condition induced greater activation in the left IFG (BA 47; Shetreet et al., in press), the weak number match condition showed decreased activation in this region.

Thus, despite some similarities between the processing of numbers and quantifiers, our results provide evidence against the classical neo-Gricean approach to numerals (e.g., van Rooij & Schulz, 2006; Horn, 1972). This approach assumes that the lexical interpretation of numerals is the weak ("at least") interpretation and that the strong ("exact") interpretation is derived by a scalar implicature that derives the strong interpretation with quantifiers. Therefore, it predicts substantially similar brain patterns for processing weak numerals and weak quantifiers, which we did not observed in the current study.

Are our results consistent with another approach proposed to explain the various possible interpretations of numerals? Here, we want to consider in particular two approaches: an enriched scalar implicature approach (Panizza & Chierchia, 2011) and a domain widening approach (based on, e.g., Breheny, 2008). According to the enriched scalar implicature approach, like in the classical neo-Gricean approach, the lexical meaning of numerals consists of the "at least" interpretation and the strong interpretation, which is derived by a scalar implicature, is the "exact" interpretation. However, this approach argues that numerals obligatorily induce scalar implicatures, whereas the implicature is optional with quantifiers. Furthermore, implicatures are assumed to be easier and faster with numerals (see, e.g., Barner & Bachrach, 2010; Huang & Snedeker, 2009a, 2009b). Thus, this approach predicts some differences in brain activations related to the implicature generation of numerals and quantifiers. For both numerals and quantifiers, it would be harder to process (and thus, induce greater activations for) the weak scalar expression (*some* and *three*) compared with the strong one (*every* and *five*). However, weak numerals should be generally easier to processes than weak quantifiers. No differences in brain activations related to implicature mismatch processing are predicted, as it is assumed that mismatch processing occurs for both numerals and quantifiers in a mismatch scenario.

The domain-widening approach assumes that numbers have the strong "exact" interpretation as their lexical meaning, unlike quantifiers (for which the "exact" interpretation, *some but not all*, is the strong interpretation). This account further assumes that the weak ("at least") interpretation is derived by a pragmatic process, which can be viewed as a sort of domain widening (for a detailed description, see Breheny, 2008). For example, *three cats are purring* says, on the "exact" (lexical) reading, that there are exactly three purring cats in indexically supplied domain (formalized as Domain D). The pragmatic process then adds some kind of quantification over this domain: For some Domain D to be determined, there are (exactly) three purring cats (which is consistent with there being more than three in some larger domain). Thus, this approach predicts that brain activations in response to numerals and quantifiers would differ on some respects. Specifically, it predicts different brain activations for the two scalar expressions in the derivation processes (of the weak interpretation for numerals and the strong interpretation for quantifiers). With regard to numerals, the weak ("at least") interpretation is predicted to be harder to generate (and thus, induce greater activations) compared with the strong interpretation. This approach also predicts that mismatch cases would be more difficult to process than the other cases for both scalar expressions.

The predictions of the two approaches concerning the brain activations of weak numerals are quite similar, and the results of our study are consistent with both of these approaches. At the same time, the increased activations for the weak numbers (even if only for the mismatched condition) in the left IFG, a region that was linked to the generation of scalar implicatures with quantifiers, may support the enriched scalar implicature approach. Whereas the domain-widening approach assumes different mechanisms for generating the pragmatic interpretation of numbers and quantifiers (domain-widening and scalar implicature, respectively), the enriched scalar implicature approach assumes that the pragmatic interpretation is derived by the same mechanism for both scalar expressions (i.e., scalar implicature). The difference according to this approach is in the ease and speed of generating the implicatures. Therefore, the enriched scalar implicature approach specifically predicts that a processing load for the weak numerals should be observed in the region linked to scalar implicature generation with quantifiers (i.e., the left IFG). The domain-widening approach does not make any specific prediction regarding the localization of the activation related to weak numerals.

Although our results show differences between numerals and quantifiers with regard to the implicature generation, the specific pattern of dissimilarities observed in the current study is not predicted by either of the above approaches. To our understanding, both approaches would predict that the weak number mismatch condition and the weak number match condition show the same pattern of brain activation in the left IFG. However, this is not what was observed in this region, where the two conditions diverged. Our experiment cannot offer a conclusive explanation for this finding. Here we can only speculate about a mechanism that could result in such activation pattern, but further research is needed to explore the differences between match and mismatch scenarios with numerals.

A possible mechanism to explain the differences between number mismatched condition and the number matched condition is feedback between the mismatch processing (attributed to the left MFG and MeFG/ACC regions) and the generation of the strong interpretation (attributed to the left IFG). In the match case, both the weak and the strong interpretations yield the same response; therefore, it may seem unnecessary to keep activating both of the interpretations. A feedback mechanism may occur, for example, if the assessment of the strong interpretation for the weak number match condition would be inhibited following the assessment of the possible responses. If, as predicted by the enriched scalar implicature approach, the numeral implicatures are generated faster (as shown in experimental results, e.g., Huang & Snedeker, 2009a, 2009b), the rapid generation may enable the mismatch processing to influence the implicature generation process and inhibit the activation in the left IFG. Under such account, the time to make the implicature with quantifiers is longer. This may weaken the influence of the mismatch processing so no attenuation of the activation in the left IFG would be observed (as shown in Shetreet et al., in press). Of course, as stated above, this account has to be verified in future research that directly explores match and mismatch contexts with numerals.

To summarize, our neuroimaging results join previous findings suggesting that numerals and quantifiers are processed differently. It specifically points to the match scenarios as the source of difference between the two scalar expressions. Thus, although we found similarities between sentences presented in the mismatch scenario, our study clearly goes against approaches that ascribe the exact same processing for numerals and quantifiers (e.g., the classical neo-Gricean approach). The activation of the weak numerals in a mismatch scenario in the left IFG (BA 47), which was linked to the processing of implicatures with quantifiers (Shetreet et al., in press), may support the enriched scalar implicature approach, although we cannot rule out other approaches to the representation and processing of numerals.

Acknowledgments

This research was supported by the Mind/Brain/Behavior Initiative at Harvard University, the International Brain Research Organization, and the Boston Children's Hospital.

Reprint requests should be sent to Einat Shetreet, Laboratories of Cognitive Neuroscience, Children's Hospital Boston, 1 Autumn St, Boston, MA 02115, or via e-mail: einat.shetreet@childrens.harvard.edu.

Notes

1. The order of the scale is based on asymmetric entailment among its members, where the stronger members entail the weaker ones. For example, on the quantifier scale, "many" entails "some" (because if many students passed the exam, then some students passed the exam) but "some" does not entail "many" (because if some students passed the exam, it does not mean that many students passed the exam).
2. Upward entailing contexts license inferences from subsets to supersets (e.g., *John bought mystery books* entails that John bought books). In downward entailing contexts, the direction of the entailment reverse, so that inferences are made from supersets to subsets (e.g., *If John bought books, it is because he had researched the author* entails that if John bought mystery books, he had researched the author).
3. As we performed the same experiment with adults and children (Shetreet, Cheirchia, & Gaab, submitted), we wanted to keep the experimental procedure identical for both groups to allow a proper comparison between them. Our experience with children indicates that they have difficulties in switching between tasks in the same run in the MRI machine. We therefore decided to present the control and experimental tasks in different runs for the adult group as well.

REFERENCES

Barner, D., & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology, 60,* 40–62.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19,* 2767–2796.

Bookheimer, S. Y. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 25,* 151–188.

Breheny, R. (2008). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics, 25,* 93–139.

Carston, R. (1998). Informativeness, relevance and scalar implicature. In R. Carston & S. Uchida (Eds.), *Relevance theory: Applications and implications* (pp. 179–236). Amsterdam: Benjamins.

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective & Behavioral Neuroscience, 7,* 367–379.

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

Chierchia, G., Fox, D., & Spector, B. (2008). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In P. Portner, C. Maienborn, & K. von Heusinger (Eds.), *Handbook of semantics* (pp. 157–168). New York: Mouton.

Dale, P., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments & Computers, 28,* 125–127.

Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron, 24,* 427–432.

Fiez, J. A. (1997). Phonology, semantics and the role of the left inferior prefrontal cortex. *Human Brain Mapping, 5,* 79–83.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. B., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2,* 189–210.

Geurts, B., Katsos, N., Cummins, C., Moons, J., & Noordman, L. (2010). Scalar quantifiers: Logic, acquisition, and processing. *Language and Cognitive Processes, 25,* 130–148.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). New York: Academic Press.

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences, 9,* 416–423.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (4th ed., pp. 819–836). Cambridge, MA: MIT Press.

Homae, F., Hashimoto, R., Nakajima, K., Miyashita, Y., & Sakai, K. L. (2002). From perception to sentence comprehension: The convergence of auditory and visual information of language in the left inferior frontal cortex. *Neuroimage, 16,* 883–900.

Horn, L. (1972). *On the semantic properties of the logical operators in English*. Doctoral Dissertation, UCLA, Los Angeles, CA.

Horn, L. (2004). Implicature. In L. R. Horn, B. G. Ward (Eds.), The Handbook of Pragmatics (pp. 3–28). Malden, MA: Blackwell.

Huang, Y., & Snedeker, J. (2009a). On-line interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology, 58,* 376–415.

Huang, Y., & Snedeker, J. (2009b). Semantic meaning and pragmatic interpretation in five-year olds: Evidence from real time spoken language comprehension. *Developmental Psychology, 45,* 1723–1739.

Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development, 2,* 77–96.

Mansouri, F. A., Tanaka, K., & Buckley, M. J. (2009). Conflict-induced behavioural adjustment: A clue to the executive functions of the prefrontal cortex. *Nature Reviews Neuroscience, 10,* 141–152.

Marty, P., Chemla, E., & Spector, B. (2013). Interpreting numerals and scalar items under memory load. *Lingua, 133,* 152–163.

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

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

Noveck, I. A., & Sperber, D. (2007). The why and how of experimental pragmatics: The case of "scalar inferences." In N. Burton-Roberts (Ed.), *Advances in pragmatics* (pp. 184–212). Basingstoke: Palgrave.

Panizza, D., & Chierchia, G. (2011). Numerals and scalar implicatures. In J. Meibauer & M. Steinbach (Eds.), *Experimental pragmatics/semantics* (pp. 129–150). Amsterdam: John Benjamins.

Panizza, D., Chierchia, G., & Clifton, C., Jr. (2009). On the role of entailment patterns and scalar implicatures in the processing of numerals. *Journal of Memory and Language, 61,* 503–518.

Panizza, D., Huang, Y. T., Chierchia, G., & Snedeker, J. (2009). Relevance of polarity for the online interpretation of scalar terms. *Proceedings of Semantics and Linguistic Theory (SALT), 19,* 360–378.

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

Sakai, K. L. (2005). Language acquisition and brain development. *Science, 310,* 815–819.

Shetreet, E., Chierchia, G., & Gaab, N. (in press). When some is not every: Dissociating scalar implicature generation and mismatch. *Human Brain Mapping*. doi:10.1002/hbm.22269.

Shetreet, E., Chierchia, G., & Gaab, N. (submitted). Linguistic inability or poor performance: Pragmatic inferences in the developing brain.

van Rooij, R., & Schulz, K. (submitted). Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy, 29,* 205–250.

Wendelken, C., Nakhabenko, D., Donohue, S. E., Carter, C. S., & Bunge, S. A. (2008). Brain is to thought as stomach is to ??: Investigating the role of rostrolateral prefrontal cortex in relational reasoning. *Journal of Cognitive Neuroscience, 20,* 682–693.

Wolfensteller, U., & von Cramon, D. Y. (2011). Strategy-effects in prefrontal cortex during learning of higher-order S-R rules. *Neuroimage, 57,* 598–607.