ARTICLE IN PRESS

Cognition xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Number as a cognitive technology: Evidence from Pirahã language and cognition ☆

Michael C. Frank [a], Daniel L. Everett [b], Evelina Fedorenko [a], Edward Gibson [a,*]

[a] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room 46-3037D, Cambridge, MA 02139, USA
[b] Department of Languages, Literatures, and Cultures, Illinois State University, Campus Box 4300, Normal, IL, 61790, USA

ARTICLE INFO

Article history:
Received 25 June 2007
Revised 3 March 2008
Accepted 27 April 2008
Available online xxxx

Keywords:
Language and thought
Number
Cross-cultural research
Indigenous peoples

ABSTRACT

Does speaking a language without number words change the way speakers of that language perceive exact quantities? The Pirahã are an Amazonian tribe who have been previously studied for their limited numerical system [Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science 306*, 496–499]. We show that the Pirahã have no linguistic method whatsoever for expressing exact quantity, not even "one." Despite this lack, when retested on the matching tasks used by Gordon, Pirahã speakers were able to perform exact matches with large numbers of objects perfectly but, as previously reported, they were inaccurate on matching tasks involving memory. These results suggest that language for exact number is a cultural invention rather than a linguistic universal, and that number words do not change our underlying representations of number but instead are a cognitive technology for keeping track of the cardinality of large sets across time, space, and changes in modality.

1. Introduction

How does language shape our understanding of number? Animals and pre-linguistic infants are able to discriminate large quantities approximately (Dehaene, 1997; Gallistel, 1990; Lipton & Spelke, 2003; Xu & Spelke, 2000) and show some understanding of exact operations with small quantities (Hauser & Carey, 2003; Wynn, 1992). However, human adults routinely manipulate exact numbers in ways that are beyond the reach of other animals even after large amounts of training (Matsuzawa, 1985; Pepperberg & Gordon, 2005). The single most important difference between the numerical cognition of humans and that of other animals is our reliance on linguistic representations of quantity – number words – to act as symbolic placeholders in complicated operations. In fact, many theorists have hypothesized that linguistic symbols play a causal role in the acquisition of exact numerical competence, allowing children to extend their abilities to reason about small numbers of objects to larger quantities (Carey, 1998; Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999).

Strong support for the view that language is involved in the acquisition of numerical competence comes from indigenous groups with limited number vocabulary in their languages. The cognition of these groups shows the hallmarks of approximate rather than exact numerical competence (Gordon, 2004; Pica, Lemer, Izard, & Dehaene, 2004). For example, recent work with the Pirahã people of Brazil (a monolingual hunter–gatherer tribe living in the Amazon rainforest) has demonstrated that the Pirahã have at best a limited inventory of words relating to number (Everett, 2005; Gordon, 2004). In addition, the Pirahã language is reported not to have singular–plural morphology,

☆ The authors MCF and EG contributed equally to this project. The authors wish to thank David Barner, Peter Gordon, Amy Perfors, Rebecca Saxe, Elizabeth Spelke, Josh Tenenbaum, Ed Vul, Nathan Witthoft, and an anonymous reviewer for their helpful comments. The first author was supported by a Jacob Javits Graduate Fellowship and the second author was supported by the EC project CHLaSC.
* Corresponding author. Tel.: +1 617 452 2474.
E-mail addresses: mcfrank@mit.edu (M.C. Frank), egibson@mit.edu (E. Gibson).

0010-0277/$ - see front matter
doi:10.1016/j.cognition.2008.04.007

meaning that there is no morphological route for representing the distinction between "one" and "many" in Pirahã (Everett, 2005). Gordon additionally found that across a variety of different tasks requiring judgments of quantity the Pirahã produced errors which grew larger systematically as quantities increased, indicating that they were probably using a strategy of approximate magnitude estimation, rather than representing numbers exactly. One particular result was especially surprising: The Pirahã made errors on a simple one-to-one matching task. In the other matching tasks, the Pirahã might have understood what was required but been unable to perform the tasks accurately; this conclusion would lead to the inference that number vocabulary is necessary for remembering large numbers accurately. However, given its lack of auxiliary cognitive demands, the failures of the Pirahã in the one-to-one matching task also suggested a potentially deeper, strong Whorfian claim: That without number words, human beings represent only approximate quantities, and that only by learning number words can humans create the concept of *exact* quantity: The idea that adding or subtracting even a single individual from a set will change the quantity of that set.

Here we investigate these two claims: The weaker claim, that language for number allows accurate memory for – and hence operations over – sets with exact cardinalities; and the stronger claim, that language for number creates the concept of exact quantity (Gelman & Gallistel, 2004; Gordon, 2004). Building on the work of Gordon (2004), we investigate both the number language (Experiment 1) and numerical abilities (Experiment 2) of the Pirahã. Consistent with previous reports, we find that the Pirahã truly have no linguistic method of expressing any exact quantity, even "one." However, despite this lack, they are able to perform exact matching tasks with large numbers of objects when these tasks do not require memory. These results militate against the strong Whorfian claim that learning number words creates the concept of exact quantity. Instead, they suggest a view of number words as a cognitive technology, a tool for creating mental representations of the exact cardinalities of sets, representations that can be remembered and communicated accurately across time, space, and changes in modality.

2. Experiment 1: Numeral elicitation

Gordon (2004) described the Pirahã language as having a numerical vocabulary corresponding to the terms "one" (*hói*), "two" (*hoí*), and "many" (*baagiso*, though he reports the variant *aibaagi*). He also noted that these terms do not have exact meanings, thus *hói* may mean "roughly one" or "small." Everett has suggested, however, that there are no numerals in the language whatsoever and that these words instead indicate "small size or amount," "somewhat larger size or amount," and "cause to come together/many" (Everett, 2005). To test these claims and establish whether Pirahã contains any absolute number terms, we simply asked Pirahã speakers to describe varying quantities of objects (roughly following the design in Pica et al., 2004).

2.1. Participants and methods

Six adult Pirahã speakers participated in the increasing elicitation condition and four participated in the decreasing elicitation condition. To elicit descriptions of quantities in the Pirahã language, we presented sets of spools of thread to our participants. In the increasing elicitation condition, we started with one spool and added spools one by one until there were 10 spools of thread. For each quantity, we asked the question "how much/many is this?" (translated into Pirahã by D.E.). In the decreasing elicitation condition, we started with 10 spools and took spools away one by one until there was only one spool remaining. The experiment was run with participants that had completed the matching tasks in Experiment 2 immediately beforehand, thus the participants were aware that we were particularly interested in the size of sets.

2.2. Results and discussion

On every trial, participants produced one of the three words *hói*, *hoí*, and *baágiso*. The proportion of each word produced for each number in the two conditions is shown in Fig. 1. In the increasing elicitation, *hói* was universally used to describe one object, *hoí* was used to describe two or more objects, and *baágiso* was used to describe quantities of three or more. These data were consistent with meanings of "one," "roughly two," and "many" for the three words. However, in the decreasing elicitation, *hói* was used to refer to quantities as large as six, *hoí* was used for quantities between 4 and 10, and *baágiso* was used for quantities between 7 and 10. Across the two tasks, none of the three words that the Pirahã produced were used consistently to refer to any particular quantity across the two tasks. Because each of the three words was used for a dramatically different range of values in the ascending and the descending elicitations, these words are much more likely to be relative or comparative terms like "few" or "fewer" than absolute terms like "one" or even proto-numbers (numerals with approximate quantities, like "roughly one," as suggested in Gordon, 2004). A proto-

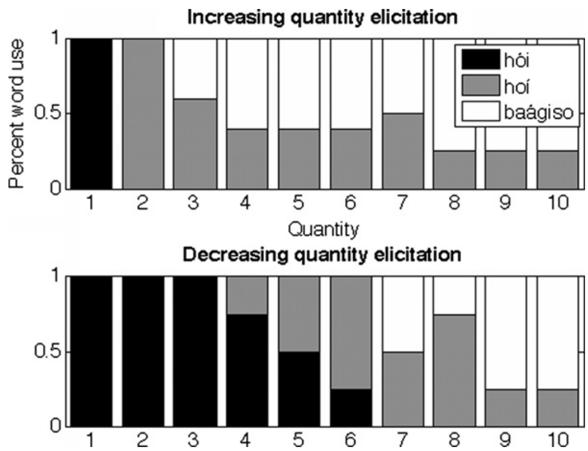

Fig. 1. Proportion of Pirahã speakers using each of the three proposed quantity words in Pirahã. Sets with different quantities were presented in increasing order and participants were asked to describe their quantity.

number referring to a fixed but approximate quantity should not change in its range of application across different contexts, and intuitively the translation "roughly one" seems misleading for a word that can be used to refer to up to six objects.

Are there other words or morphemes indicating exact number in Pirahã? We give two arguments against this possibility. First, no other numerals have been reported by Everett, Gordon, Keren Madora, Steve Sheldon, or Arlo Heinrichs, researchers that have collectively been working with the Pirahã for more than 50 years. Second, no other words or morphological markers were produced with any consistency in our experiment, meaning that if there were a word or morpheme for "exactly one" in Pirahã it was not elicited in nine independent viewings of a single object (in the case of a single word or morpheme that we failed to recognize). Thus if such a word or morpheme exists it is at best extremely low frequency and rarely used in discussions of quantity. While we cannot rule out this possibility, it appears unlikely.

Whereas many languages have only a limited vocabulary of number words (Menninger, 1969), we do not know of any other language in which this type of elicitation has been performed. Thus, to our knowledge Pirahã is the first case in which a language has been documented as lacking any linguistic device for expressing the quantity "one." However, assessing how rare this property is will require experiments like the elicitation we performed to be carried out with a substantial sample of the many other languages with restricted numeral systems.

3. Experiment 2: Matching tasks

In order to assess the numerical cognition of the Pirahã, we performed a series of matching tasks similar to those used by Gordon (2004). Our intent was to make a systematic test of Pirahã speakers' abilities in exact numerical tasks with varying perceptual and memory demands. In his studies, Gordon found a decrease in performance as quantities increased across a wide variety of tasks. These results were consistent with the use of an analog magnitude estimation strategy in every task, suggesting that the Pirahã might have fundamentally different representations of large numbers than speakers of languages with a recursive count list. In addition, the results implied that the Pirahã did not appreciate the difference between two large numbers of approximately but not exactly equal quantities (e.g., 7 and 8) and hence might lack even the notion of *exact quantity*. Although Gordon's results were suggestive in this direction, they were conducted with a small sample of participants (only four individuals, all male, provided data for many of the experiments), without a translator, and with varying procedures between experiments. Thus, we attempted to replicate his results with a larger sample and a more systematic procedure.

3.1. Participants and methods

Fourteen adult Pirahã speakers (seven men and seven women, the majority of the adult population of one village) participated in the hidden, uneven, orthogonal, and one-to-one matching tasks and nine of those individuals participated in the nuts-in-a-can task. The materials for the tasks – spools of thread and uninflated rubber balloons – were chosen both because the Pirahã were already familiar with them and because they were small and easy to manipulate. All participants performed five tasks (except for five participants who did not perform the "nuts-in-a-can" task), in the following order: A one-to-one match task, an uneven match task, an orthogonal match task, a hidden match task, and a "nuts-in-a-can" task. In each trial of each task, the experimenter presented some quantity of spools and then asked the participant to put out the same quantity of balloons in a line. This continuity of response across all five tasks (which were always performed during a single experimental session lasting not more than 30 min) ensured that failure in the more difficult tasks was not due to changes of response format. In the one-to-one task, the experimenter placed an evenly-spaced line of spools on the table and the participant was asked to put out a matching line of balloons. In the uneven-match task, the experimenter grouped the spools into irregular sets of two, three, or four spools within the line. In the orthogonal match task, the experimenter placed an evenly-spaced line of spools on the table stretching away from the participant, orthogonal to the matching line of balloons. In the hidden match task, the experimenter placed the spools in a line and then concealed them behind an opaque folder. Finally, in the nuts-in-a-can task, the experimenter dropped the spools one by one into an opaque cup into which the participant could not see.

In order to make sure that the Pirahã participants understood our tasks, we first modeled each task for each participant (with the exception of the uneven match tasks, which was judged to be very similar to the one-to-one match task), with one experimenter (E.G.) testing a second experimenter (M.C.F.) on the quantities two and three. In modeling the one-to-one and uneven matching tasks, balloons were placed immediately in front of the spools of thread (suggesting direct correspondence). We then asked the participant to respond on the quantities two and three, repeating each trial with correction in the case of any errors. Although these trials were not explicitly labeled as training trials, in cases of confusion or error they helped to clarify the requirements of the task. These two measures together helped to ensure that the Pirahã did not perform poorly due to misunderstandings. No participants made any errors on the training trials for one-to-one match task; five participants each made a single error on the hidden match task and one other participant required multiple corrections; two participants each made a single error on the orthogonal matching task; and two participants each made multiple errors on the nuts-in-a-can task.

3.2. Results and discussion

The performance of Pirahã participants is plotted in Fig. 2. Consistent with the results reported by Gordon (2004), performance on the orthogonal match, hidden match and "nuts-in-a-can" tasks decreased as quantity increased. For quantities of four and above, the standard deviation appeared constant relative to the quantity being estimated,

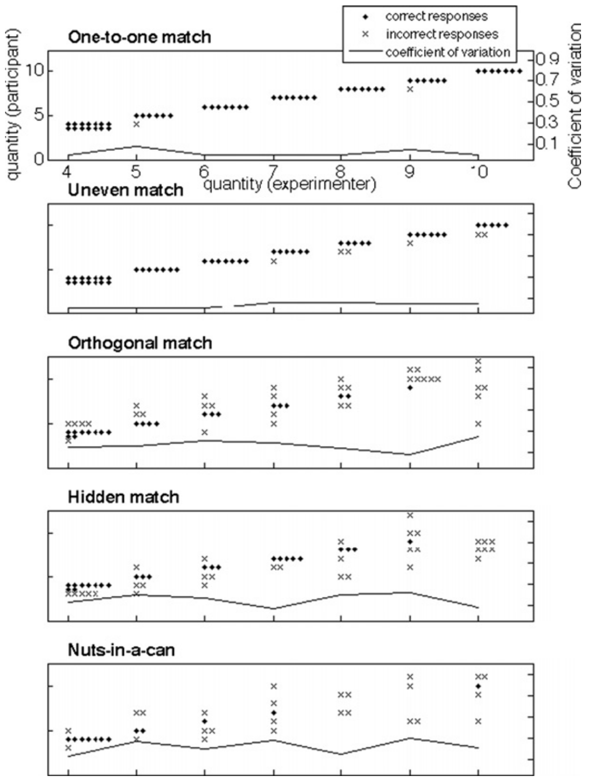

Fig. 2. Performance and coefficient of variation plotted by task. The left-hand axes plot quantity of spools provided by the experimenter on the *X*-axis and quantity of balloons matched to the spools by Pirahã participants on the *Y*-axis. Correct responses are marked with a dot, while incorrect responses are marked with an *x*. Multiple correct responses at a given quantity are staggered. The right-hand axes plot the coefficient of variation at each quantity.

congruent with Weber's law (a signature of analog magnitude estimation, see e.g., Whalen, Gallistel, & Gelman, 1999). The mean coefficient of variation (standard deviation/mean) was 0.16 for the orthogonal match, 0.15 for the hidden match task and 0.21 for the "nuts-in-a-can" task (plotted for each quantity on the right-hand axis of each graph); these figures are highly comparable to the aggregate coefficient of variation of 0.15 for quantities of 4 and above reported by Gordon (2004).

However, performance on the one-to-one matching task was nearly perfect, and performance on the uneven match task was close to ceiling as well. Of 14 participants, only a single participant made any errors on the one-to-one matching task (a total of 54 of 56 trials were performed correctly); we observed 6 errors total in the uneven match task (50 of 56 trials correct, with 10 of 14 participants making no errors). Thus, performance as measured by participants' percent correct responses in the uneven match was lower than performance in the one-to-one match, but not significantly so (paired $t(13) = 1.30$, $p = .21$).[1] In contrast, participants' performance on the one-to-one match differed significantly from performance in the orthogonal match (24/56 trials correct, $t(13) = 5.95$, $p < .001$),[2] hidden match (24/56 trials correct, $t(13) = 6.51$, $p < .001$),[3] and "nuts-in-a-can" (12/36 trials correct, $t(8) = 9.71$, $p < .001$)[4] tasks. The orthogonal match, hidden match, and "nuts-in-a-can" tasks did not differ significantly from one another (all values of t less than .40, with all values of $p > .70$).[5] Results were comparable in their level of significance when these analyses were performed across items rather than participants.

While our results on the more difficult of the two tasks largely replicate those of Gordon (2004), the performance of our participants in the one-to-one and uneven matching tasks were qualitatively different; however, we suspect that theoretically unimportant aspects of the testing materials and environment may have caused the differences in performance (P. Gordon, personal communication). In particular, Gordon's participants were tested with AA batteries on an uneven surface, which may have led the objects to move around inadvertently within a trial. In contrast, our tests were conducted with spools of thread (placed on their flat side) and balloons on a flat table in an enclosed hut. The objects did not move within a trial unless the participants moved them and there were no outside distractions. Furthermore, although it is possible that the presence of training trials may have contributed to the lack of errors in this task, it seems unlikely that the errors Gordon observed were due to a lack of such training trials. In particular, as Gordon argued, the errors he observed in the one-to-one and uneven matching tasks increased with the quantity of the set (indicating a source of error in matches, like rolling batteries) rather than appearing randomly (indicating a subset of participants who simply did not understand the task). Thus, we find it more likely that it was the circumstances of testing, rather than participants' understanding of the tasks, that contributed to the differences between our results and those of Gordon (2004).

More generally, we suggest that the pattern of performance on the three tasks we conducted roughly reflects the demands of each task. The one-to-one and uneven matching tasks require no memory for the exact cardinality of a set (whatever it may be), only an understanding that the set *has* an exact quantity and thus that one item more or less than the quantity of the set to be matched is not a correct response. The perfect or near-perfect performance of the Pirahã in these tasks indicates that they were able to appreciate the necessity of matching the quantity of objects exactly rather than approximately. In contrast, the orthogonal match, the hidden match, and the "nuts-in-a-can" task all required not only understanding that an exact match was necessary, but also transferring information about the cardinality of sets across time and space using an exact rather than approximate representation. Though the transfer across space in the orthogonal match task – a task in which participants showed highly similar perfor-

[1] Because t-tests may not be appropriate for means over categorical data (since they are not normally distributed), we also give the results of Wilcoxon signed rank tests (a non-parametric test equivalent to a paired t-test) which in all cases confirmed the results of the parametric tests. For this comparison, $p = 0.38$.

[2] Nonparametric $p < 0.001$.

[3] Nonparametric $p < 0.001$.

[4] Nonparametric $p = 0.004$.

[5] Nonparametric p values all >.70.

mance to the hidden match and "nuts-in-a-can" tasks – may seem trivial to numerically-sophisticated, it also requires a strategy either for maintaining the exact correspondence of objects (via the use of eye-movements or fingers as placeholders to compare items one by one) or for representing the cardinality of the set to be matched (via the use of verbal representations). Because the Pirahã had no linguistic tools for verbal encoding and no practice in using non-verbal strategies, they were unable to perform these tasks, leading to errors of increasing magnitude and frequency with increases in the cardinality of the set to be remembered.

4. Discussion

A total lack of exact quantity language did not prevent the Pirahã from accurately performing a task which relied on the exact numerical equivalence of large sets. This evidence argues against the strong Whorfian claim that language for number creates the concept of exact quantity (and correspondingly, that without language for number, any task requiring an exact match would be impossible). Instead, the case of Pirahã suggests that languages that can express large, exact cardinalities have a more modest effect on the cognition of their speakers: They allow the speakers to remember and compare information about cardinalities accurately across space, time, and changes in modality. Visual and auditory short-term memory are highly limited in their capacity and temporal extent (Baddeley, 1987). However, the use of a discrete, symbolic encoding to represent complex and noisy perceptual stimuli allows speakers to remember or align quantity information with much higher accuracy than they can by using their sensory short-term memory. Thus, numbers may be better thought of as an invention: A cognitive technology for representing, storing, and manipulating the exact cardinalities of sets.

Do the Pirahã then possess mental representations of the cardinalities of large sets? We do not believe that our experiments show evidence supporting this hypothesis. Success in the one-to-one and uneven match tasks requires participants to understand that the addition or subtraction of exactly one object makes a match incorrect, even for large quantities. Thus, the Pirahã understand the concept of one (in spite of having no word for the concept). Additionally, they appear to understand that adding or subtracting one from a set will change the quantity of that set, though the generality of this knowledge is difficult to assess without the ability to label sets of arbitrary cardinality using number words. However, the one-to-one matching task itself can be completed via a simple algorithm, "put one balloon down next to one spool." At no point during the task must participants represent the cardinality of the entire set. They need only to understand that, in the application of this algorithm it is *exactly* one balloon that must be matched to *exactly* one spool. Thus, our experiments support the hypothesis that the concept of exact quantity is not created by language, while suggesting that the ability to remember the cardinalities of large sets is enabled by learning number words.

Where does this leave the Whorf hypothesis, the claim that speakers of different languages see the world in radically different ways? Our results do not support the strongest Whorfian claim. However, they are consistent with several recent results in the domains of color (Gilbert, Regier, Kay, & Ivry, 2006; Uchikawa & Shinoda, 1996; Winawer et al., 2007) and navigation (Hermer-Vazquez, Spelke, & Katsnelson, 1999). In each of these domains, language appears to add a second, preferred route for encoding and processing information. In the case of color, language enables faster performance in search, better discrimination, and better memory when target colors can be distinguished from distractors by a term in the participant's language. However, verbal interference – which presumably blocks access to linguistic routes for encoding – eliminates this gain in performance, suggesting that the underlying perceptual representations remain unmodified. Likewise in the case of navigation: The use of particular linguistic devices allows (though does not require, see e.g., Li & Gleitman, 2002) efficient compressive navigational strategies. But again, under verbal interference these strategies are not accessible and participants navigate using strategies available to infants and non-human animals.

In both of these domains, as well as in our work on number, language plays a fundamentally compressive role, allowing the efficient encoding of information about quantity, color, and spatial orientation. However, in cases where the appropriate code is suppressed or not useful, speakers perform in the same way as speakers of languages that do not even possess the relevant vocabulary. The color, number, and navigational vocabularies of different languages thus do not seem to alter the underlying cognitive or perceptual processes of speakers of those languages directly. Instead, like other technologies such as alphabetic writing (O'Connor, 1996), languages give their users a new route for the efficient encoding of experience.

References

Baddeley, A. D. (1987). *Working memory*. Oxford, UK: Oxford University Press.

Carey, S. (1998). Knowledge of number: Its evolution and ontogeny. *Science, 282*, 641–642.

Dehaene, S. (1997). *The number sense*. New York: Oxford University Press.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science, 284*, 970.

Everett, D. L. (2005). Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology, 46*, 621–646.

Gallistel, C. R. (1990). *The organization of learning Cambridge*. MA: MIT Press.

Gelman, R., & Gallistel, C. R. (2004). Language and the origin of numerical concepts. *Science, 306*, 441–443.

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences, 103*, 489–494.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science, 306*, 496–499.

Hauser, M. D., & Carey, S. (2003). Spontaneous representations of small numbers of objects by rhesus macaques: Examinations of content and format. *Cognitive Psychology, 47*, 367–401.

Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology, 39*, 3–36.

Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition, 83*, 265–294.

Please cite this article in press as: Frank, M. C. et al., Number as a cognitive technology: Evidence from Pirahã language ..., *Cognition* (2008), doi:10.1016/j.cognition.2008.04.007

Lipton, J. S., & Spelke, E. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science, 14*, 396–401.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature, 315*, 57–59.

Menninger, K. (1969). *Number words and number symbols: A cultural history of numbers*. Cambridge, MA: MIT Press.

O'Connor, M. (1996). The alphabet as a technology. In P. T. Daniels & W. Bright (Eds.), *The world's writing systems* (pp. 787–794). New York: Oxford University Press.

Pepperberg, I. M., & Gordon, J. D. (2005). Number comprehension by a grey parrot (*Psittacus erithacus*), including a zero-like concept. *Journal of Comparative Psychology, 119*, 197–209.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*, 499–503.

Uchikawa, K., & Shinoda, H. (1996). Influence of basic color categories on color memory discrimination. *Color research and application, 21*, 430–439.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10*, 130–137.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104*, 7780–7785.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358*, 749–750.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1–B11.