Language as a cognitive technology: English-speakers match like Pirahã when you don't let them count

Michael C. Frank, Evelina Fedorenko, and Edward Gibson
{mcfrank, evelina9, egibson}@mit.edu
Department of Brain and Cognitive Sciences, 43 Vassar Street
Cambridge, MA 02139 USA

Abstract

The Pirahã, an Amazonian hunter-gatherer tribe, lack words for numbers and are unable to complete simple matching tasks when the tasks require memory for exact quantities (Gordon, 2004; Frank et al., in press). Here we show that American participants perform similarly to the Pirahã when asked to execute the same kinds of matching tasks under verbal interference. These results provide support for the hypothesis that number words act as a "cognitive technology": a method for quickly and efficiently storing information via abstraction. We review a variety of other evidence supporting this proposal from the domains of color, navigation, and theory of mind.

Keywords: Numerical cognition; verbal interference; language and thought; Whorf hypothesis; Pirahã.

Introduction

How does knowing a language affect the way you are able to perceive, act and reason in the world? Do the differences between languages cause systematic differences in the cognition of their speakers? These questions about the relationship between language and thought have been among the most controversial in cognitive science for many years. However, recent evidence from non-linguistic and cross-cultural populations has given some insight into this relationship in domains such as number, color, navigation, and theory of mind.

In this paper, we propose a unifying account of these strands of evidence, which we call the "cognitive technology" hypothesis (Dascal, 2002; Frank et al., in press):

> *Rather than altering underlying representations, languages instead help their speakers accomplish difficult or intractable cognitive tasks by providing abstractions which allow for the efficient storage and processing of information.*

This hypothesis synthesizes a number of existing theoretical ideas. Vygotsky (1986) suggested that language could be a scaffold for action by providing external support for difficult tasks. Similarly, several authors have suggested that language works to code experience and that Whorfian effects stem from differences in the way experience is coded (Hunt & Agnoli, 1991), or that language augments cognition by providing an external resource for information storage and transformation (Clark, 1998). While these proposals are deeply related to our own, we believe that the contribution of the view here is to synthesize these ideas into an account in which the role of language as a tool for abstraction explains a number of recently identified "Whorfian" effects, while still acknowledging the existence of core cognitive abilities which are unaltered by language (Spelke & Kinzler, 2007; Carruthers, 2002).

To support this hypothesis, we first discuss numerical cognition, a case study of language and conceptual change (Carey & Spelke, 1994) which has been studied across a variety of populations. In particular, we focus on the finding that the Pirahã—an Amazonian hunter-gatherer group who lack words for exact numbers, are able to perform simple one-to-one matching tasks but unable to perform numerical tasks that require remembering exact quantities (Everett, 2005; Frank et al., in press; Gordon, 2004). Here we present experimental data showing the same pattern of performance in American participants who performed these matching tasks while simultaneously engaged in verbal shadowing. We then discuss these data in light of results from several other domains.

Numerical cognition in infants and non-human primates is thought to be subserved by two distinct systems (Feigenson et al., 2004). The parallel-individuation ("object file") system is related to visual attention and object tracking and is used to track the identity of small numbers of discrete objects. In contrast, the analog magnitude system is used to represent large, approximate quantities and can operate over arbitrarily large quantities. Parallel individuation is precise but only functions for quantities below three or four; in contrast, analog magnitude estimation exhibits a constant coefficient of variation (error relative to the size of the set being estimated) (Whalen et al., 1999).

The relationship between a lack of number language and reliance on these core, pre-linguistic numerical systems was first documented via case studies of two Amazonian groups, the Pirahã (Gordon, 2004) and the Mundurukú (Pica et al., 2004). The Pirahã presented a particularly interesting case: their language was reported to have words roughly corresponding to the concepts of "one," "two," and "many", and they were unable to perform a variety of simple matching tasks. Gordon interpreted these results as providing evidence for a strong Whorfian claim: that without language for number, the Pirahã had no notion of exact quantity and were thus unable even to put objects in one-to-one correspondence.

In recent work (Frank et al., in press), we provided evidence that Pirahã in fact has no words for exact quantities whatsoever; the previously reported numerical terms appear to be comparative or relative terms. In addition, we showed that—contra Gordon (2004)—the Pirahã succeeded in simple one-to-one matching tasks, suggesting that the concept of exact quantity (the idea that adding or subtracting one object makes a difference) does not depend on linguistic knowledge. However, we replicated Gordon's results that the Pirahã had considerable difficulty in performing matching tasks which required memory for exact quantities.

Recent research by Trick (2005) supports the conclusion that aspects of number are dependent on language. They showed that counting accuracy was differentially impaired by a complex verbal interference task (saying the letters F' and S' in alternation) compared with a simple verbal task (saying the letter S') and both simple and complex tapping tasks. However, since this work tests performance in an explicitly verbal task under verbal interference, it does not directly address the hypothesis that numerical cognition relies on language more generally, not simply when a verbal response is needed.

In the current study, we provide the first direct test of the hypothesis that memory for exact quantities, but not one-to-one matching, is dependent on language. We performed the same set of matching tasks that we used with the Pirahã with English-speaking participants in Boston, MA. Pilot testing revealed that these tasks were trivial for our English-speaking participants; they were able to complete all the matching tasks with no errors simply by counting. Thus, in order to investigate the role of linguistic knowledge in this task, we introduced a verbal interference task that participants performed concurrently with the matching tasks. Verbal interference tasks are typically used to prevent the rehearsal of verbal information in short-term memory. As our interference task we chose verbal shadowing (repeating meaningful speech immediately after hearing it) (Hermer-Vazquez et al., 1999; Newton & de Villiers, 2007). We hypothesized that, absent the ability to count, the performance of English-speaking participants in these tasks would have the same qualitative signatures as the performance of the Pirahã.

Experimental Data

Methods

Participants We recruited 20 participants from MIT and the surrounding community; our participants varied in age from 18 to 50, approximately matching the distribution of ages in our Pirahã population (though exact matching was of course impossible because the Pirahã could not report their own exact ages). Participants received $10 as compensation.

Procedure Participants were first familiarized with the verbal shadowing task: they were instructed to listen to short clips from the Radio News Corpus (Ostendorf et al., 1995) and to repeat the words spoken by the announcer as quickly as possible. After their performance was judged to be fluent by the experimenter and they reported that they were comfortable with the task, they were given instructions for the matching tasks.

Each participant completed five matching tasks (as described in Frank et al., in press), in the following order: a one-to-one matching task, an uneven matching task, an orthogonal matching task, a hidden matching task, and a nuts-in-a-can task. Each matching task required the participant to observe some quantity of spools of thread and to put out a line of un-inflated balloons exactly matching the quantity of spools that they saw (these items were chosen because they were the same stimulus items we used with the Pirahã). For each task, participants were tested once on quantities of spools from four to twelve (in a random order) and the number of balloons they put out was recorded by the experimenter. No feedback was given.

In the one-to-one and uneven matching trials, the experimenter placed the spools one by one in a line running from the participant's left to their right. The spools were evenly spaced in the one-to-one task and broken into smaller groups of one to four in the uneven task. The line of balloons placed by participants was parallel to the line placed by the experimenter, so the participants could simply put balloons in one-to-one correspondence with spools to succeed in the task.

The orthogonal matching task was identical to the one-to-one task except that the line of spools ran from close to the participant to further away, rather than from left to right. The hidden match task was identical to the one-to-one task except that the line was hidden by the experimenter after the spools were placed (by placing a manila folder in front of the spools). Finally, in the "nuts-in-a-can task," the experimenter placed spools one by one into an opaque cup.

On each trial, the experimenter would begin by starting the audio (which the participant listened to over headphones). Once the participant had begun shadowing, the experimenter placed the spools one by one in the task configuration. Once the experimenter had finished, the participant began placing balloons to indicate quantity. When finished placing balloons, the participant indicated that the trial was finished by pressing a key to end the audio.

Results

Participants' individual data is shown in Figure 1 along with the data we collected with the Pirahã. Additional analyses are shown in Table 1. For the English speakers as for the Pirahã, the one-to-one and uneven matching tasks were easiest, followed by the orthogonal and hidden matching tasks; the nuts-in-a-can task was by far the hardest. Thus, to a first approximation, occupying the verbal resources of speakers of a language with numbers produced a pattern of data remarkably similar to the data of speakers of a language without numbers. These data confirm our hypothesis: while performance in one-to-one matching tasks was largely preserved without the use of number language, performance on more memory-intensive tasks was dramatically impaired.

However, there were some differences between the Pirahã and our English speaking participants. We compared the English data to the Pirahã data via an ANOVA, predicting mean performance across participants with language group and matching task as factors. Both factors and their interaction were highly significant predictors in the model ($F(1,70) = 33.61, p < .001$ for language group, $F(4,70) = 33.76, p < .001$ for task, and $F(4,70) = 5.30, p < .001$ for their interaction), suggesting that the English speakers were

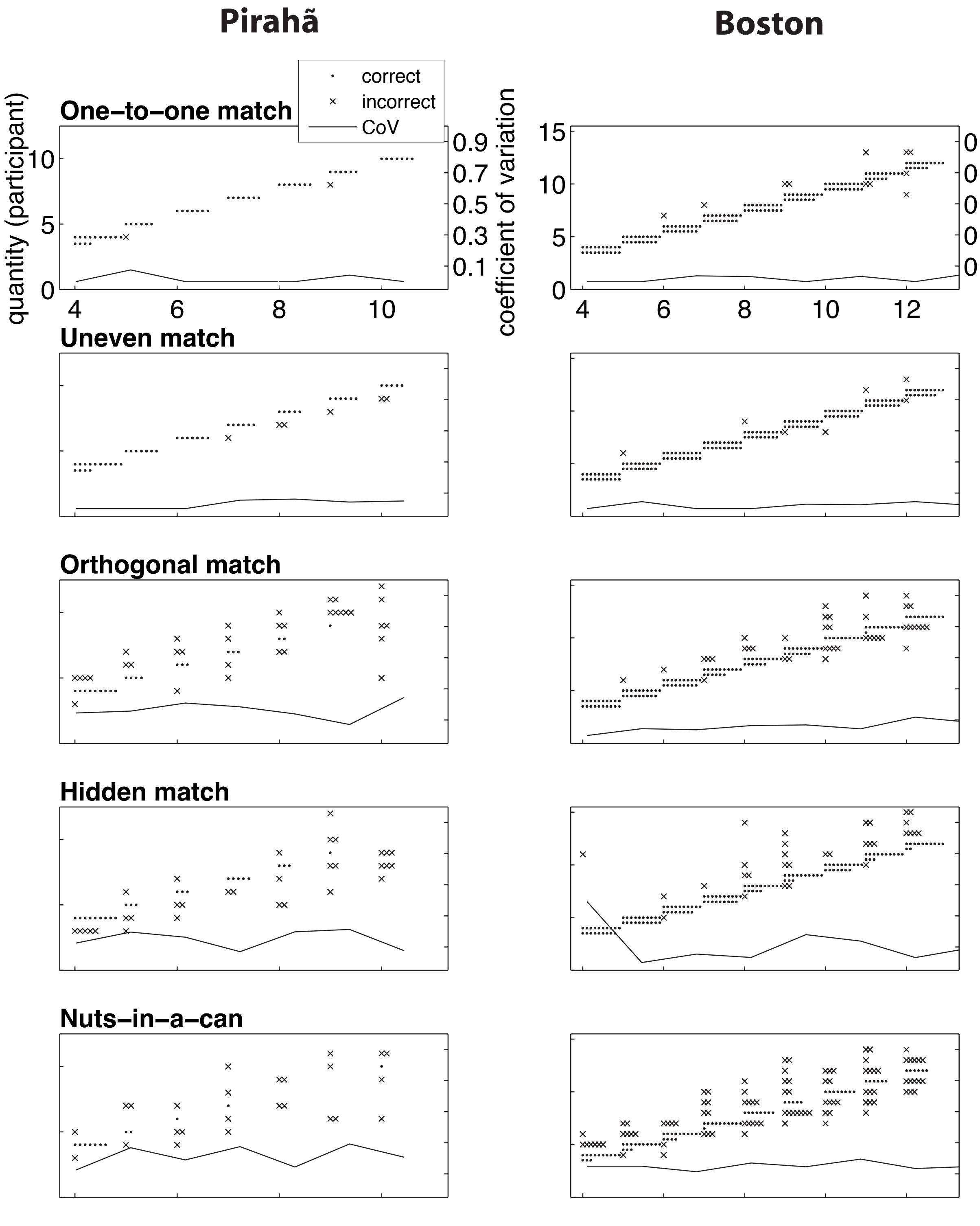

Figure 1: Comparison of data collected from the Pirahã (from Frank, Everett, Fedorenko, & Gibson, in press) and data from the verbal shadowing experiment carried out with English-speaking participants in Boston. Left-hand axes show the match between the quantity of balloons placed by the participant (Y axis) and the quantity of spools placed by the experimenter (X axis). Each datapoint is an individual trial for an individual participant. Xs mark incorrect trials while dots mark correct trials; marks are jittered to reflect the quantity of trials at each quantity for each group. The right hand axes measure the coefficient of variation (mean / standard deviation), as shown for each quantity by the solid line.

		Percent correct	Mean error	Mean COV	COV slope	COV r^2	COV p-value
English	One-to-one	0.94	0.03	0.02	0.006	0.39	0.07
	Uneven	0.96	0.03	0.02	0.003	0.15	0.31
	Orthogonal	0.78	0.11	0.06	0.010	0.65	0.01
	Hidden	0.82	0.29	0.11	-0.014	0.10	0.41
	Nuts-in-a-can	0.47	0.23	0.14	-0.003	0.06	0.53
Pirahã	One-to-one	0.96	0.04	0.02	-0.002	0.03	0.71
	Uneven	0.89	0.12	0.03	0.010	0.64	0.03
	Orthogonal	0.39	0.50	0.16	0.002	0.01	0.86
	Hidden	0.43	0.58	0.15	-0.003	0.01	0.83
	Nuts-in-a-can	0.32	0.39	0.21	0.009	0.08	0.53

Table 1: Statistical comparison of results between Pirahã and Boston participants. Percent correct refers to the total proportion of trials of a particular type that were correct. Mean error is the average difference between the mean of participants' responses and the target value (small values indicate that participants' estimates were correct across the group, despite being variable for any individual trial). Mean COV refers to the mean coefficient of variance (mean divided by standard deviation) for a task; a smaller COV indicates more precise estimates. COV slope is the slope of the best linear fit to the coefficient of variance for that trial. COV r^2 and p-value are the r^2 and p-values assigned to the linear trend in the coefficient of variance (larger r^2 values indicate growth in the coefficient of variance as the quantities being estimated grow, indicating that analog magnitude estimation may not be responsible for the errors.

overall better at the tasks, but that the differences between groups were not constant across tasks.

Examining tasks individually, the one-to-one and uneven matching tasks were extremely easy for both groups and the number of errors was relatively small, suggesting that participants in both groups were able to make inferences about exact correspondence. In addition, those errors that participants did make in these tasks were most likely not the results of analog magnitude estimation. Analog magnitude estimation is typically characterized by a constant coefficient of variance (mean divided by standard deviation) (Gordon, 2004; Whalen et al., 1999). For the Pirahã we found a significant linear slope in the coefficient of variation in the uneven match; and for the English-speakers, we found a trend towards a linear slope in the one-to-one matching task. These patterns could have been created instead by adding a constant probability of error to matching each item (a process which, in contrast to magnitude estimation, would create an increasing COV). While not conclusive, these correlations suggest a different source of error in these tasks from the error inherent in magnitude estimation.

For the hidden and orthogonal match tasks, the Pirahã and the English-speakers may have been using different strategies. While these two tasks were more difficult for both groups, the Pirahã performed far worse than the English speakers. We speculate that English-speakers were able to use a variety of strategies to mitigate the difficulty of these tasks. This conclusion is supported, following the same logic as in the matching tasks above, by the significant linear trend in the English speakers' orthogonal match COV.

Finally, the "nuts-in-a-can" task was the purest example of a task that was extremely difficult without language for number. There was no hint of a linear trend in the coefficient of variation data for either group, suggesting that both groups were using the analog magnitude system to perform this task.

Discussion

We asked English-speakers to perform a variety of matching tasks while performing verbal interference. They were unable to count during verbal interference and thus they were forced to rely on other strategies to perform these tasks. In the simplest one-to-one and uneven matching tasks, participants were relatively accurate, indicating that these tasks do not depend significantly on verbal resources (a finding confirmed by the success of the Pirahã on these tasks as well). In contrast, in the nuts-in-a-can task, both the Pirahã and the English speakers were highly impaired (and their responses showed the signature of analog magnitude estimation), suggesting that this task was highly dependent on verbal resources. The orthogonal match and hidden matching tasks fell somewhere in between. The similarity between the performance of the two groups lend support to the claim that the concept of "exact match" does not depend on language in either its genesis (as shown by the Pirahã data, c.f. Frank et al., in press) or its use (as shown by the current data). However, both learning and using the ability to remember exact quantities larger than three or four appears to depend crucially on verbal mechanisms. Further comparative evidence for these claims comes from Nicaraguan signers who had incomplete or non-existent knowledge of the recursive count list and showed a similar pattern of impairments (Flaherty & Senghas, 2007).

It is worth noting that our study has several methodological limitations. First and most significantly, MIT undergraduates and Boston natives are not well-matched controls for Pirahã. In addition to knowing how to count, our Boston participants also have a lifetime of formal education in a culture radically different from that of the Pirahã with all of the cognitive differences implied by this different background (Scribner & Cole, 1973). We are currently working to rectify this issue by identifying groups with similar educational circumstances but knowledge of number words.

In addition, there remains the possibility that the additional cognitive load imposed by the secondary task might have been responsible for our results, rather than the verbal component of the task specifically. Future research should address this limitation via the use of a matched control task (as in Hermer-Vazquez, Spelke, & Katsnelson, 1999 and Newton & de Villiers, 2007). However, previous research on this topic has found differential effects of complex verbal interference relative to non-verbal interference tasks (Trick, 2005), thus it is plausible that the verbal component of our shadowing task was directly responsible for the results we observed.

General Discussion

Taken together with the Pirahã data, our current results suggest a clear picture of the relationship between language and numerical cognition. All human beings (and many other species) share a variety of core numerical capacities. However, languages which contain recursive count lists allow their speakers to transcend these capacities and attain genuinely new numerical abilities. These novel abilities depend both on having a new representation (number words representing exact cardinalities) and on the abilities of speakers to learn new operations over this representation (addition, division, factorization, etc.). The resulting capacities can be as simple as remembering exact quantities over time or as complex as long division or calculus.

As demonstrated by our current results, however, the addition of this new level of representation does not substantially alter the initial core numerical abilities, which are still accessible and on which numerically-savvy speakers can still rely when their verbal resources are otherwise occupied. Put another way, when it exists, language for number represents a preferred route for processing numerical information, but in the absence or inaccessibility of this route, the original core abilities of object individuation and magnitude estimation are still present and accessible. Thus, language is—in the case of number—a cognitive technology, as in our original statement of the hypothesis. We speculate that the cognitive technology hypothesis holds in several other domains as well. In the remainder of this paper, we briefly review evidence supporting this claim from the domains of color, navigation, and theory of mind.

Languages vary considerably in the ways that they represent information about color (Kay et al., 2003). While some languages contain only a small selection of color words, others have much more elaborate vocabularies. A large literature has dealt with the Whorfian question of whether the presence of linguistic boundaries produce differences in color perception. Although this topic is still controversial, recent psychophysical evidence suggests a new consensus that there are relatively small but reliable effects of language on color memory and color discrimination; crucially these effects are greater in the visual hemi-field corresponding to the left hemisphere of the brain (specialized for language in right-handers) and are removed by verbal interference but not by comparable spatial interference (Gilbert et al., 2006; Winawer et al., 2007). Though more controversial, the cross-linguistic evidence collected by Roberson and colleagues (Roberson & Henley, 2007) also supports this view. Thus, the evidence on color can be summarized in a parallel fashion to the evidence on number. Even for simple perceptual discriminations, color words provide a second route for processing, increasing efficiency; when this route is unavailable, speakers fall back on their perceptual abilities (which remain unchanged by their linguistic knowledge).

In the domain of spatial navigation as well, languages encode information about navigation using a variety of different devices (Levinson, 2003). While both rats and young children who have not mastered linguistic encoding for spatial navigation have been shown to use only geometric information about the proportions of a room to reorient themselves, human adults are able to use multiple sources of information (Cheng, 1986; Hermer & Spelke, 1994, 1996). Hermer-Vazquez et al. (1999) tested adults in reorientation tasks using both verbal and non-verbal interference tasks and found that verbal, but not non-verbal, interference dramatically impaired the integration of landmark and geometric information. When denied access to verbal resources, participants relied instead on a purely geometric strategy.

Finally, reasoning about others' beliefs and desires ("theory of mind") may also rely crucially on linguistic knowledge. Recent research by Newton & de Villiers (2007) suggests that verbal interference (but not a matched non-verbal task) disrupts performance in a non-verbal false-belief task. This work follows a body of developmental evidence suggesting links between particular aspects of linguistic development and the ability to pass tasks which rely on the representation of others' beliefs (de Villiers & de Villiers, 2000) and more recent research suggesting that populations with access to limited syntactic resources in their language, such as speakers of early versions of Nicaraguan Sign Language, may also be impaired in their false-belief reasoning (Pyers, in press).

Thus, evidence from four domains—number, color, navigation, and theory of mind—suggests a middle path between many of the traditional extreme positions in the Whorfian debate. Abstractions of these types arise differentially across languages due to the particular communicative and cognitive demands of their speakers, as in the evolution of functionally-specific specified number systems (Beller & Bender, 2008) and recent evidence suggesting that color terms represent efficient partitions of perceptual space (Regier et al., 2007). However, languages do not appear to alter the original representations underlying complex cognitive tasks. Instead, they enable their speakers to perform complex tasks (often differentially across languages) by providing abstractions for efficient information processing.

Acknowledgments

The first author was supported by a Jacob Javits Grant. The authors would like especially to thank I. Shauni Deshmukh for help in data collection. Additional thanks are due to Dan Everett, Rebecca Saxe, Nathan Witthoft, and Ed Vul.

References

Beller, S., & Bender, A. (2008). The limits of counting: Numerical cognition between evolution and culture. *Science*, *319*(213), 213-215.

Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (p. 169-200). Cambridge, UK: Cambridge University Press.

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences*, *25*(06), 657-726.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, *23*, 149-178.

Clark, A. (1998). Magic words: how language augments human computation. Cambridge University Press.

Dascal, M. (2002). Language as a cognitive technology. *International Journal of Cognition and Technology*, *1*(1), 33-59.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind.* Hove, UK: Psychology Press.

Everett, D. L. (2005). Cultural constraints on grammar and cognition in pirahã. *Current Anthropology*, *46*(4), 621-646.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307-314.

Flaherty, M., & Senghas, A. (2007). Numerosity and number signs in deaf nicaraguan adults. *Proceedings of the 31st Annual Boston University Conference on Language Development*.

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (in press). Number as a cognitive technology: Evidence from pirahã language and cognition. *Cognition*.

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*(2), 489-494.

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

Hermer, L., & Spelke, E. (1994). A geometric process for spatial reorientation in young children. *Nature*, *370*, 57-59.

Hermer, L., & Spelke, E. (1996). Modularity and development: the case of spatial reorientation. *Cognition*, *61*, 195-232.

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*(1), 3-36.

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review*, *98*(3), 377–389.

Kay, P., Berlin, B., Maffi, L., & Merrifield, W. R. (2003). *The world color survey*. Palo Alto, CA: CSLI Press.

Levinson, S. C. (2003). *Space in language and cognition*. Cambridge, UK: Cambridge University Press.

Newton, A. M., & de Villiers, J. G. (2007). Thinking while talking: Adults fail nonverbal false-belief reasoning. *Psychological Science*, *18*(7), 574-579.

Ostendorf, M., Price, P. J., & Shattuck-Hufnagel, S. (1995). *The boston university radio news corpus* (Tech. Rep.). Boston University.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(15), 499-503.

Pyers, J. (in press). Constructing the social mind: Language and false-belief understanding. In S. C. Levinson & N. Enfield (Eds.), *The roots of human sociality.* New York: Berg Publishers.

Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, *104*(4), 1436-1441.

Roberson, D., & Henley, J. R. (2007). Color vision: Color categories vary with language after all. *Current Biology*, *17*(15), R605-R607.

Scribner, S., & Cole, M. (1973). Cognitive consequences of formal and informal education. *Science*, *182*(4112), 553-559.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89-96.

Trick, L. M. (2005). The role of working memory in spatial enumeration: Patterns of selective interference in subitizing and counting. *Psychonomic Bulletin and Review*, *12*(4), 675-681.

Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: MIT Press.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science*, *10*(2), 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*(19), 7780-7785.