

Linguistic relativity and numeric cognition: New light on a prominent test case

CALEB EVERETT¹

University of Miami

Introduction

The past two decades have seen both a resurgence and transformation of research on linguistic relativity. This is due in large part to the influential work of researchers such as Lucy (1992a, 1992b) and Levinson (1997, 2003, *inter alia*). Research on linguistic relativity has morphed from the Whorfian (1956), often anecdotally based enterprise, into one that is oriented around experiments of various sorts conducted among speakers of different languages. These languages are frequently the topic of inquiry because of some specific grammatical features they contain, features that potentially affect speakers' cognitive orientations vis-à-vis a given semantic domain. Research of this type has resulted in evidence for language-mediated or influenced thought in a wide array of nonlinguistic tasks related to numeric cognition (De Cruz & Pica 2008, Pica et al. 2004), gender perception (Konishi 1993, Flaherty 2001), spatial and directional construal (Levinson 2003), substance classification (Lucy & Gaskins 2001, Imai & Mazuka 2007), the perception of time (Boroditsky et al. 2011, Boroditsky 2001), and even the perception of colors (Gilbert et al. 2006, Drivonikou et al. 2007).

In their review of recent work on linguistic relativity, Wolff & Holmes (2010:1) make the following observation:

While we do not find support for the idea that language determines the basic categories of thought or that it overwrites preexisting conceptual distinctions, we do find support for the proposal that language can make some distinctions difficult to avoid, as well as for the proposal that language can augment certain types of thinking.

¹ The author wishes to thank those Pirahã who participated in this study. He is particularly grateful as well to Keren Madora, who translated for him during research among the people, and who discussed many of the ideas presented in this work.

In this paper I will focus on a particular cognitive domain, namely numeric representation. For the data considered, pertaining to an indigenous Amazonian language well known to most linguists (see e.g. D. Everett 2009), we will see that there is considerable support for the notion that language can augment certain kinds of thinking. In particular, a certain linguistic feature, viz. number terminology, can serve as a ‘conceptual tool’ (Wolff & Holmes 2010) that augments a certain kind of thinking, numeric cognition. The data suggest that when speakers lack number terminology they struggle with basic quantity recognition tasks, and therefore that number terminology augments in a critical fashion numeric thinking.

1 Linguistic effects on numeric cognition: Potential test cases

Work related to numeric cognition and the linguistic relativity hypothesis has generally focused on languages with very modest, or completely absent, systems for expressing cardinal numerosities (De Cruz & Pica 2008). For example, the results in Pica et al. (2004) suggest that speakers of Mundurukú tend to struggle with tasks that require precise representation of numerosities greater than three, a fact that is most plausibly motivated by the paucity of number terms in that language.

As Hammarström (2010) notes, there is only a handful of languages that can truly be considered anumeric. Perhaps the most well-documented case is that of Pirahã, a language spoken in southwest Amazonia that lacks any precise number words (D. Everett 2005). In that language, *hoi* signifies “small size or amount,” *hoí* indicates “somewhat larger size or amount,” and *baágiso* means to “cause to come together” or “many.” The imprecision of these terms is demonstrated experimentally in Frank et al. (2008). In other less well-documented cases anumericity has been claimed, though experimental work is still required to buttress such claims. For example Xilixana is another South American language that is said to lack all number words, including ‘one’ (Hammarström 2010). Another Amazonian language that is claimed to lack native number words altogether is Jarawara (Dixon 2004), a member of the small Arawá family. However, follow-up work has suggested conclusively that native number words do exist in Jarawara (C. Everett, under review), and in all well-documented Arawá languages.

Given that the absence of numerals in Pirahã is now so clearly documented, this case is arguably crucial to our understanding of the potential effects of anumeric language on numeric cognition. To date, three extensive experimental studies have been undertaken in an attempt to better understand the role of number terms on the basic recognition, recall, and manipulation of quantities. Below I synthesize some of the major findings from these studies, including my own. I also present new data on very recent work among the people. I will claim that the results so far obtained among these people are consistent with the notion that

number terminology can serve as a ‘conceptual tool’, and that the data present a clear example in which language serves a fundamental role in augmenting certain cognitive abilities, in this case the ability to exactly recognize quantities exceeding three. To that extent, these data can be added to the growing literature providing evidence for linguistic effects on facets of non-linguistic cognition.

2 A summary of results across three studies

Frank et al. (2008) conducted two word-elicitation tasks that corroborated D. Everett’s (2005) claims that the three aforementioned number-like terms cannot actually be considered number words, at least not ones denoting precise numerosities. In one task, Pirahã speakers were presented first with one spool of thread, and asked to provide a number term for the quantity provided. In every case they used the word *hoí*. The researchers then added spools of thread to the presented array iteratively, and after each spool was added asked the participants to identify the new quantity. This task was termed the ‘increasing quantity’ elicitation task. For this task the speakers did use *hoí* in all cases in which two spools of thread were presented. However, they also used this term to refer to as many as seven spools of thread. They employed *baágiso* to refer to quantities ranging from three to ten spools. In the second task, Pirahã speakers were presented first with ten spools of thread, and asked to name this quantity. In most cases they provided *baágiso*, though they also used *hoí* in some cases. The researchers then subtracted spools of thread iteratively, and after each spool was subtracted asked the participants to identify the new quantity. *Baágiso* was used for quantities ranging from seven to ten spools, *hoí* was used for four to ten spools, while *hói* was utilized for one to six items. These findings support Everett’s (2005) claim that these words are not precise number terms, and suggest that Pirahã may be the most anumeric language documented in the literature.

Gordon (2004) performed a series of quantity recognition tasks among the speakers of two villages. These tasks included a basic one-to-one recognition task, an orthogonal matching task, and a brief-presentation/hidden matching task. For the one-to-one matching task, the Pirahã were individually presented with an evenly-spaced line of objects and asked to produce a matching line of objects of a different type parallel to the presented line. For the hidden matching task, stimuli were presented and shortly thereafter concealed. For the orthogonal matching task, a line of stimuli was placed in front of each participant, perpendicular to the line then produced by the participants. The coefficient of variation (standard deviation of responses divided by mean for each set size) of the Pirahã responses generally hovered at 0.15 for all tasks. This figure is generally suggestive of analog-estimation strategies on the part of the Pirahã matching the quantities (Weber’s law).

Frank et al. (2008) replicated Gordon’s one-to-one matching, orthogonal

matching, and hidden matching experiments among the Pirahã. They employed the same general methods used by Gordon (2004). For each of the tasks, they presented the Pirahã with an array of spools of thread, and the participants were asked to match the array with a line of empty rubber balloons. The Pirahã were tested individually, and in all cases the tasks were modeled for them by the researchers prior to individual trials. In general, their results were consistent with Gordon's (2004)—the people's performance generally deteriorated as the quantity in question was increased. Crucially, however, the authors failed to replicate the most extreme finding in Gordon (2004), specifically that the Pirahã struggle with the mere recognition of exact quantities greater than three. In fact, the Pirahã they tested were generally quite adept at matching a presented array of spools with an equal array of balloons when no spatial re-orientation or mental recollection of the stimuli was required.

Given the importance of the Pirahã case to the discussion of linguistic effects on numeric cognition, and given the crucial difference between the previous sets of results, I helped perform a series of tasks among a group of speakers not tested in either previous study. Three of these tasks were exact methodological replications of the tasks described in Frank et al. (2008), based as well on the three tasks from Gordon (2004) described above. A few of the tasks were quite different in that they employed cross-modal stimuli. The latter tasks were relatively modest in scope, and are described in 3. The former were more extensive, and the results obtained for them are described in detail in C. Everett & Madora (in press). Next I summarize the relevant data for these three replication tasks.

For all three tasks, the speakers were presented with uniformly-spaced lines of spools of thread, and asked to match those lines with equal lines of rubber balloons. These objects are familiar to the people, having been used as trade goods previously. Most crucially these objects were chosen since they were utilized in Frank et al. (2008) and C. Everett & Madora (in press) sought to replicate exactly the findings of that study. Just as in Gordon (2004), for the basic matching task the stimuli were presented on a table in front of the seated participants, parallel to the edge of the table. In the case of the orthogonal matching task, the spools were placed orthogonally to the edge of the table (in line with the participants' sagittal plane), and the speakers were asked to match the quantity of spools in a straight line parallel to the edge. For the hidden matching task, a line of spools of thread was presented parallel to the edge of the table, and after several seconds the line was covered by a sheet of cardboard. The participants were asked to place a matching line of balloons on the opposite side of the cardboard, parallel to the presented line of spools.

For the sake of greater methodological clarity, in (1) I provide a picture of a correctly-matched line of stimuli. This picture represents one of 56 documented trials for this task.

Linguistic relativity and numeric cognition

(1) Example of basic one-to-one matching task. Correct response for a trial with eight target stimuli.



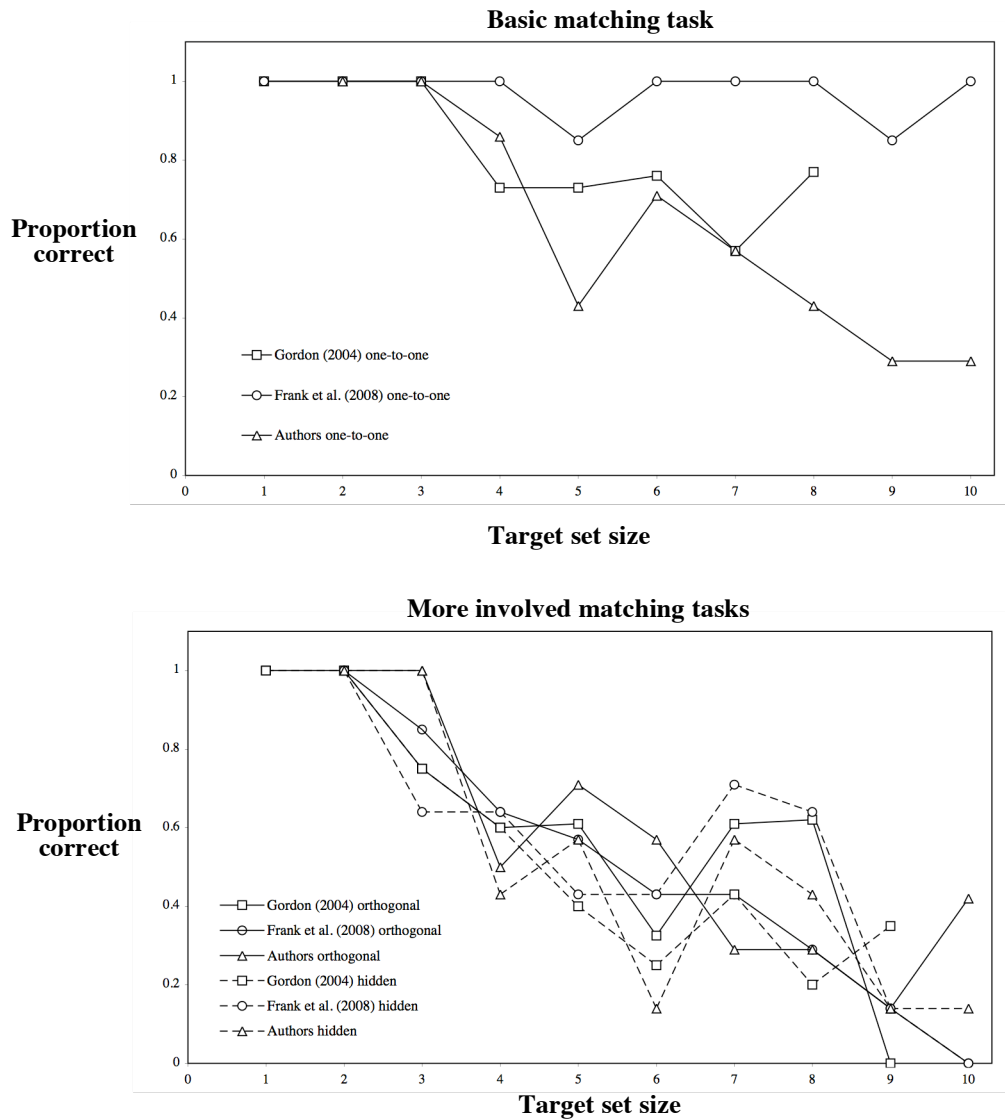
Seven Pirahã adults participated in Gordon (2004). A total of fourteen adults participated in Frank et al. (2008), and fourteen different speakers participated in C. Everett & Madora (in press). For the orthogonal matching task, 24/56 trials in C. Everett & Madora (in press) contained correct responses in which the Pirahã matched the presented stimuli with an array equal in number. This is the identical ratio of correct responses for that task in Frank et al. (2008). For the hidden matching task, 24/56 trials in C. Everett & Madora (in press) and Frank et al. (2008) contained correct responses. For the one-to-one matching task, however, 54/56 trials in Frank et al. (2008) contained correct responses, while only 32/56 did in C. Everett & Madora (in press). When individuals' proportions of correct responses were contrasted, the difference between the results of the two studies were found to be highly significant in the case of the one-to-one matching task ($t(13), p = 0.000$). When this metric was used to contrast the Pirahãs' responses across the two other tasks, the differences across studies were not found to be significant ($p > .05$ in each case).

The coefficient of variation for all the tasks in Gordon (2004), C. Everett & Madora (in press) hovered around 0.15, consistent with the use of analog estimation by the people (rather than task incomprehension). This coefficient was also obtained in Frank et al. (2008), with the exception of the basic matching task.

In short, the results in C. Everett & Madora (in press) for the orthogonal-matching and hidden-matching tasks are very similar to those in Frank et al. (2008). They are also similar to those in Gordon (2004), as evidenced by similar proportions of correct responses calculated according to set size. This is apparent in the second box of Figure (2). The results in C. Everett & Madora (in press) for the basic matching task, which did not involve recall or spatial manipulation, are

similar to those in Gordon (2004), but not Frank et al. (2008) as is apparent in the first box of Figure (2). In general, the results for all three studies suggest that the speakers of this anumeric language struggle with the recognition of exact correspondences between numerosities over three. In 4 we offer an explanation of the disparate results in Frank et al. (2008) vis-à-vis the basic matching task. First, though, we provide additional findings recently gathered among the people.

(2) Proportions of correct responses for various matching tasks. (Taken from C. Everett & Madora, in press.)



3 New cross-modal data

One criticism that could be made of the previous studies of Pirahã numeric cognition, including C. Everett & Madora (in press), is that they rely exclusively on data of a specific kind. For all studies the Pirahã speakers were asked to perform a visual-tactile task that does not appeal to the auditory modality. While the results of the studies are consistent with the implementation of analog estimation during the tasks, rather than task-comprehension failure, one wonders whether the Pirahãs' performance might benefit from a greater array of cross-modal tasks. To begin exploring this issue, I conducted two brief tasks with ten Pirahã speakers. Only a modest number of trials were conducted for each of these tasks, due to a limited window for research with the Pirahã in question.² Nevertheless, the results of these tasks are worth discussing here since they yield further support for the suggestions in C. Everett & Madora (in press) and Gordon (2004) that the Pirahã struggle with simple quantity recognition.

The two tasks I conducted involved physical actions, and in one case auditory stimuli. The tasks were: stomping-with-log repetition and rowing action repetition. In the case of both tasks, the Pirahã participants were asked to repeat as closely as possible the actions performed by myself. For the stomping-with-log repetition task, I created a series of booming noises by simultaneously stomping my right foot and smashing a narrow log on the ground. The participant was then asked to repeat the action. The task was first modeled between myself and another non-Pirahã, in an attempt to make it clear that the objective of the task was to imitate the number of stomping actions. While this task may seem somewhat esoteric, it was chosen because it employs auditory and kinesthetic information and relates to a behavior common to Pirahã culture. During the 'dance' that takes place throughout the night during full moons, the people have often been observed to stomp out a series of noises, with their foot and with a narrow log simultaneously, while walking in a circle. Given their familiarity with this motion, and given the somewhat rhythmic nature of the stomping that occurs during the 'dance', it seemed natural to utilize this motion experimentally.

The second task also involved a behavior that is familiar to all Pirahã, rowing with a paddle. For the task, I created a series of rowing motions with a paddle, alternating from side-to-side. The participant was then asked to mimic my actions. The entire task was once again modeled by myself and another non-Pirahã, until it seemed clear that the number of rowing actions was of interest. A pause was made between each rowing action, ostensibly so that the actions would be perceived as discrete units. This task was also selected because it represents a behavior that is common to Pirahã culture. The Pirahã, who are uniformly excellent rowers, are often observed paddling with a series of symmetrical motions.

² Access to the speakers of this language is generally quite limited for a variety of reasons. In the case of these cross-modal tasks, I was limited to one-day for the experiments.

Given this fact, I imagined that the frequently symmetrical nature of canoe rowing might facilitate their ability to recall and store the number of actions witnessed.

I suspected that, because these two tasks involved behaviors that were comparatively common to Pirahã culture, when contrasted to those in C. Everett & Madora (in press), Gordon (2004), and Frank et al. (2008), and because they involved actions of a different modality, their number recall for the tasks might show improvement. Based on the few dozen trials I have so far conducted, these suspicions are far from confirmed. In fact, the performance of the Pirahã on these trials suggests that they face even greater difficulties with these sorts of tasks, perhaps since they involve some recollection or perhaps because the request made of them is so unfamiliar (despite the familiarity of the action).

For both tasks, all ten speakers tested presented the correct number of actions only in the case of the number one. That is, if one stomping or rowing action was produced by myself, one stomping or rowing action was then produced by the participant. For numbers greater than two, however, incorrect responses outnumbered correct responses in all cases. The responses for the rowing-action task contained enough errors to suggest task incomprehension, but only for numerosities greater than two.

For the stomping-action task, I tested numbers from 1-5. The means of the answers, according to each target size (number of stomping-actions produced by experimenter), were as follows: 1:1, 2:3.75, 3:4.25, 4:4.75, and 5:6.5. In other words, the magnitude of errors was generally quite high for numbers 2-5. For instance, when five stomping actions were produced by myself, the average number of response actions was 6.5. As the number of target actions increased, however, so did the number of participant actions. This suggests that the participants did recognize an increase in the quantity of actions and attempted to match the quantity, albeit imprecisely.

These cross-modal data, while modest in scope, suggest strongly that speakers of Pirahã struggle with recognizing or recalling the exact number of actions witnessed. It is worth mentioning that I have also utilized other tasks involving a variety of other actions (e.g. clapping) with a smaller set of speakers, and have yet to observe anything that would lead me to suspect that some other cross-modal task might exist for which the Pirahã would demonstrate heightened quantity recognition.³ In short, the data so far collected among the speakers of this anumeric language suggest that they struggle with exact quantity recognition. The only exception to this trend in the experimental data is the finding in Frank et al. (2008) *vis-à-vis* simple one-to-one matching. We are naturally left to wonder what might account for the disparate findings in that study.

One possibility is that the findings in Frank et al. (2008) were due to greater clarity on the part of the experimenters, i.e. that the Pirahã tested in the remaining studies were confused by the tasks in a way that those tested in Frank

³ Such observations are consistent with the fact that many previous attempts at teaching math skills among the people have failed. (D. Everett 2005)

et al. (2008) were not. Given that the other studies relied on a fluent translator, given that their results were characterized by relatively constant coefficients of variation (inconsistent with task incomprehension), and given that the results across control tasks for all studies are so similar, this possibility is highly improbable, however. One plausible account of this disparity requires some elaboration of the setting of the research conducted for Frank et al. (2008), which is provided next.

4 Variation across villages?

The approximately 700 Pirahã are dispersed over numerous villages along the Maicí. These villages are typically small, most often with a dozen or so adults. The territory of the people stretches south from the mouth of the river to the point at which the river is crossed by a federal highway (BR 230), which is actually a dirt road through the jungle. There are several villages within 20 km of this road. Two of these were used as locations for the research in Gordon (2004) and a third was used in our own research. The latter village is located at 7°48' S, 62°20' W, and is nearly adjacent to a bridge crossing the Maici river.⁴ The data presented in Frank et al. (2008) were based on research in another location, the Xagiopai village. Xagiopai is located about 50 km from BR 230, at 7°21' S, 62°16' W.

All of the Pirahã remain almost exclusively monolingual despite contact with Brazilians for over two centuries. (See D. Everett 2005.) In the case of the Pirahã living near BR 230, this contact is primarily with transient Brazilians. The case at Xagiopai is much different, however. FUNAI (the Brazilian indigenous agency) and FUNASA (the Brazilian health organization) have maintained relatively extensive operations in the village for over a decade. The Xagiopai village is the only Pirahã village with such prominent government presence, and is also the village in which an SIL linguist, Keren Madora, resided during the early-to-mid 2000's. Unbeknownst to the authors of Frank et al. (2008), one of the primary foci of Madora was to teach the Pirahã how to count. To that end, she employed numerous quantity recognition tasks, of the sort she helped develop for Gordon (2004:496). Crucially, she also introduced various numeric neologisms into the language. This was the first time this had been done in the language. According to Madora, the performance of the Pirahã improved if they learned these neologisms.

Given these facts, it is less surprising that the Pirahã at Xagiopai did much better on the one-to-one match documented in Frank et al. (2008), when contrasted with those documented in the other relevant studies. Admittedly I cannot be certain that the disparate performance in Frank et al. (2008) was due to the neologisms coined by Madora. Nevertheless, at present I believe that this is the most plausible interpretation of the data (see C. Everett & Madora in press).

⁴ The experiments took place in a Brazilian-owned house dozens of meters off the reservation.

5 Discussion and conclusion

Studies such as Wynn (1992), Lipton & Spelke (2003), and Xu & Spelke (2000) suggest that pre-linguistic infants are capable of exactly recognizing exact quantities less than three, and are also capable of approximating larger quantities. These abilities appear to be based on two distinct neurophysiological systems, as evidenced in Dehaene et al. (1999) and Lemer et al. (2003), *inter alia*. Carey (2001) and others have suggested that number words serve a crucial ontogenetic purpose, namely to conjoin these two core systems. The Pirahã data are consistent with these results from the developmental literature.

The absence of number words in Pirahã is, according to D. Everett (2005, 2009), the result of general cultural constraints that result in a series of other typologically-remarkable features in this language. Were the Pirahã familiarized with such words as children they would undoubtedly excel at the tasks described here. The people have clearly been reluctant to borrow such words or systematically incorporate numeric neologisms. This reluctance stands in stark contrast to other cultures with modest number systems, which have typically adopted number terms from other languages. It is important to stress that the Pirahã excel in their environment, and that they show relatively little interest in the very acquisition the ‘tool’ of number terminology, much as they have little interest in most tools and artifacts offered by outsiders for which, according to them, they would have little use (D. Everett 2005). In short, while I believe the data discussed here are consistent with relativistic effects, it is important to recognize that these effects could arguably be due, ultimately, to a more general cultural factor, namely the opposition to the incorporation of number words into their language.

The data discussed here are consistent with the notion that speakers of an anumeric language lack a ‘conceptual tool’, a series of number words, which is nearly universal to all cultures. The lack of this linguistic/conceptual tool apparently results in strong cognitive effects when the quantity-recognition abilities of speakers of such a language are contrasted with the abilities of the speakers of a numeric language. More generally, since a particular feature of the Pirahã language, namely anumericity, apparently has demonstrable effects on non-linguistic cognition, namely number recognition, the data discussed above add to the growing literature on the linguistic relativity hypothesis.

References

- Boroditsky, L. 2001. Does language shape thought? English and Mandarin speakers' conceptions of time. *Cognitive Psychology* 43.1-22.
- Boroditsky, L., Fuhrman, O., & McCormick, K. 2011. Do English and Mandarin speakers think differently about time? *Cognition* 118.123-129.
- Carey, S. 2001. Cognitive foundations of arithmetic: Evolution and ontogenesis.

Linguistic relativity and numeric cognition

- Mind and Language*, 16(5), 37-55.
- De Cruz, H., & Pica, P. 2008. Knowledge of number and knowledge of grammar: Number as a test case for the role of language in cognition. *Philosophical Psychology* 21.437-441.
- Dehaene, S., E. Spelke, P. Pinel, R. Stanescu, & S. Tsivkin. 1999. Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science* 284.970-974.
- Dixon, R.M.W. 2004. *The Jarawara language of southern Amazonia*. Oxford University Press.
- Drivonikou, G.V., P. Kay, T. Regier, R.B. Ivry, A.L. Gilbert, A. Franklin, & I.R.L. Davies. 2007. Further evidence that Whorfian effects are stronger in the right hemisphere than the left. *Proceedings of the National Academy of Sciences* 104.1097-1102.
- Everett, C. Under review. A closer look at a supposedly anumeric language.
- Everett, C. & K. Madora. In press. Quantity recognition among speakers of an anumeric language. *Cognitive Science*.
- Everett, D. L. 2005. Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology* 46.621-646.
- Everett, D.L. 2009. Pirahã culture and grammar: A response to some criticisms. *Language* 85.405-442.
- Feigenson, L., S. Dehaene, & E. Spelke. 2004. Core systems of number. *Trends in Cognitive Sciences* 8.307-314.
- Flaherty, M. 2001. How a language gender system creeps in perception. *Journal of Cross-cultural Psychology* 32.18-31.
- Frank, M., D. Everett, E. Fedorenko & E. Gibson. 2008. Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition* 108.819-824.
- Gilbert, A., T. Regier, P. Kay & R.B. Ivry. 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. Numeric cognition without words: Evidence from Amazonia. *Science* 306.496-499.
- Hammarström, H. 2010. Rarities in numeral systems. In Wohlgemuth, Jan, & Michael Cysouw (eds.), *Rethinking universals: How rarities affect linguistic theory*, 11-60. Berlin: De Gruyter Mouton.
- Imai, M. & R. Mazuka. 2007. Language-relative construal of individuation constrained by universal ontology: Revisiting language universals and linguistic relativity. *Cognitive Science* 31.385-413.
- Konishi, T. 1993. The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research* 22.519-534.
- Lemer, C., S. Dehaene, E. Spelke, & L. Cohen. 2003. Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia* 41.1942-1958.

Caleb Everett

- Levinson, Stephen. 1997. Language and cognition: The cognitive consequences of spatial description in Guugu Yimithirr. *Journal of Linguistic Anthropology* 7.98-131.
- Levinson, S. 2003. *Space in language and cognition*. Cambridge: Cambridge University Press.
- Lipton, J. S., & E. Spelke. 2003. Origins of number sense: Large-number discrimination in human infants. *Psychological Science* 14.396-401.
- Lucy, J. 1992a. *Language diversity and thought*. Cambridge: Cambridge University Press.
- Lucy, J. 1992b. *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
- Lucy, J. & Gaskins, S. 2001. Grammatical categories and classification preferences. In Bowerman, Melissa & Stephen Levinson (eds.), *Language Acquisition and Conceptual Development*, 257-283. Cambridge: Cambridge University Press.
- Pica, P., C. Lemer, V. Izard & S. Dehaene. 2004. Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306.499-503.
- Whorf, B. 1956. *Language, Thought, and Reality: Selected Writings*. Carroll, J. (ed.) Cambridge, Massachusetts: MIT Press.
- Wolff, P. & Holmes, K. 2010. Linguistic relativity. *WIREs Cognitive Science*.
- Wynn, K. 1992. Addition and subtraction by human infants. *Nature* 358.749-750.
- Xu, F., & E. Spelke. 2000. Large number discrimination in 6-month-old infants. *Cognition* 74.B1-B11.

Caleb Everett
University of Miami
Department of Anthropology
PO Box 248106
Coral Gables, Florida 33124-2005

caleb@miami.edu