COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 36 (2012) 130–141
Copyright © 2011 Cognitive Science Society, Inc. All rights reserved.
ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/j.1551-6709.2011.01209.x

Quantity Recognition Among Speakers of an Anumeric Language

Caleb Everett,[a] Keren Madora

[a]*Department of Anthropology, University of Miami*

Received 24 September 2010; received in revised form 28 March 2011; accepted 11 April 2011

Abstract

Recent research has suggested that the Pirahã, an Amazonian tribe with a number-less language, are able to match quantities > 3 if the matching task does not require recall or spatial transposition. This finding contravenes previous work among the Pirahã. In this study, we re-tested the Pirahãs' performance in the crucial one-to-one matching task utilized in the two previous studies on their numerical cognition, as well as in control tasks requiring recall and mental transposition. We also conducted a novel quantity recognition task. Speakers were unable to consistently match quantities > 3, even when no recall or transposition was involved. We provide a plausible motivation for the disparate results previously obtained among the Pirahã. Our findings are consistent with the suggestion that the exact recognition of quantities > 3 requires number terminology.

Keywords: Number; Numerical cognition; Linguistic relativity; Pirahã

1. Introduction

Perhaps more than any of the attested cases of languages with simple number systems (cf. De Cruz & Pica, 2008; Hammarström, 2010), Pirahã is remarkably incapable of denoting exact quantities. As Everett (2005) notes, there are three number-like words in the language, but these terms are imprecise. *Hói* indicates a ''small size or amount,'' *hoí* indicates ''somewhat larger size or amount,'' and *baágiso* means to ''cause to come together'' as well as ''many.'' Frank, Everett, Fedorenko, and Gibson (2008) performed a series of word-elicitation tasks demonstrating that these number-like words are not used to refer to specific numerosities and that the language is indeed anumeric.

Correspondence should be sent to Caleb Everett, Department of Anthropology, University of Miami, PO Box 248106, Coral Gables, FL 33124–2005a. E-mail: caleb@miami.edu

Gordon (2004) presented remarkable results on quantity recognition among the Pirahã. Most significantly, Gordon suggested that Pirahã speakers were unable to accurately perform exact one-to-one matching between more than three objects. Instead the Pirahã employed analog estimation strategies in order to discriminate such quantities. Frank et al. (2008, p. 820) note that the most remarkable finding in Gordon's study is that the Pirahã made errors on a simple one-to-one matching task that did not require auxiliary spatial manipulation or recall. This finding suggests that Pirahã speakers lack the basic one-to-one correspondence principle so crucial to numeric cognition (see, e.g., Carey, 2001, p. 40), and more generally that number words may be required for exact recognition of quantities > 3.

A number of studies on numerical cognition, including Wynn (1992), Xu and Spelke (2000), Lipton and Spelke (2003), and Feigenson, Dehaene, and Spelke (2004), have demonstrated that pre-linguistic infants as well as other species are capable of representing 1–3 exactly and are also capable of approximating numerical magnitudes. Such findings are consistent with the existence of two core representational systems: a small exact number system and a large approximate number system. Research on the ontogeny of numerical cognition (e.g., Sarnecka & Gelman, 2004) suggests that number words and integer lists serve a crucial developmental role, although the nature of this role is a matter of some debate (see discussion in Condry & Spelke, 2008, p. 23). In particular, Condry and Spelke's (2008) results suggest that these words allow for the extension of the aforementioned representational systems and for the creation of a ligature between them. This ligature putatively enables the exact recognition of larger numerosities. Taken in concert with such findings, Gordon's (2004) results suggest that individuals without access to such terms lack a means of conjoining the systems, as evidenced by their inability to exactly recognize one-to-one correspondences for quantities exceeding 3. The results therefore support Carey's (2001, p. 39) claim that ''the capacity to represent positive integers is a cultural construction that transcends core knowledge.''

Frank et al. (2008) conducted a number of tasks similar to those in Gordon (2004). In the case of those tasks requiring additional cognitive demands, for example, a task requiring the orthogonal matching of objects, the results of the study generally replicated those in Gordon. Their results suggested that the Pirahã employed analog estimation when attempting to recall quantities or mentally transpose them across space. Significantly, however, the researchers failed to replicate the finding in Gordon vis-à-vis the Pirahãs' performance on a simple one-to-one matching task that did not require recall or transposition. Instead the researchers found that Pirahã speakers were generally quite adept at the task in question. Their findings suggest that number words serve as a ''cognitive technology for representing, storing, and manipulating the exact cardinalities of sets'' (823) but imply that this technology is not required for the mere recognition and representation of positive integers > 3.

2. Experiment 1: Replication of matching tasks

The disparity between the Pirahãs' performances for the basic one-to-one matching task, across the two relevant studies, is puzzling. Frank et al. (2008) suggest that the

poor performance of the Pirahã on the relevant task in Gordon (2004) was due to a methodological flaw in his study. For his one-to-one matching task, subjects were presented with a line of AA batteries and asked to create a parallel line of the same number of AA batteries. In some cases rolling batteries may have deleteriously affected the Pirahãs' performance. Keren Madora (K.M.), who served as translator during periods of Gordon's initial work among the Pirahã, and who assisted in the design of some tasks utilized in that study (p. 496), doubts this factor alone could account for the gross discrepancies between the results of the one-to-one matching tasks in the two studies. She notes that, for preliminary tasks not reported in Gordon, various stimuli besides batteries were employed with similar results. This claim is supported by video data recording during the general period of Gordon's research, which demonstrates inaccurate performance on the one-to-one matching task without any concomitant rolling of stimuli. We conducted another study on Pirahã numerical cognition in order to help resolve the crucial discrepancy between the two previous studies.

2.1. Participants and methods

The exact methods and materials employed in the one-to-one matching task, hidden matching task, and orthogonal matching task described in Frank et al. (2008) were utilized here. The latter two tasks served as control experiments. There were 14 subjects (8 females), all the adults of one village. None of the subjects had participated in either of the two previous studies in question. For the one-to-one matching task, the Pirahã were individually presented with an evenly spaced line of spools of thread, each several centimeters apart in accordance with the previous studies, and asked to present a matching line of empty rubber balloons parallel to the line of spools of thread. All instructions were carried out in Pirahã, and participants were specifically requested to make a line that was equal to that provided. No overt feedback was given during the tasks, in keeping with the previous studies. For the hidden matching task, the spools of thread were presented and shortly thereafter concealed behind a sheet of cardboard. For the orthogonal matching task, the experimenters placed a line of evenly spaced spools stretching away from each participant, perpendicular to the line of balloons matched by the participant. As in Frank et al., the authors modeled the tasks for each Pirahã individual, with one experimenter testing the other with first two and then three spools of thread. Each participant was then presented with two spools of thread and then three spools of thread, and asked to match each quantity. These trials served to test their comprehension of the task. No errors were made by any participants on these initial trials, for any of the tasks. Subjects were then presented with lines of spools ranging in number from 4 to 10. To prevent fatigue effects for individual tasks, not all numbers from 4 to 10 were presented to each participant, and 56 trials were run for each of the three tasks—as in Frank et al. We ensured that all participants were tested for an equal number of smaller and larger quantities, presented in randomized order. Sessions lasted approximately 25 min and were documented via pictures and video.

2.2. Results

Figure 1 depicts the proportions of correct responses, for each task. The results from our study and the two previous studies are contrasted. The proportion of correct responses generally drops significantly for numbers exceeding 2 or 3. This is true for all documented tasks, with the exception of the one-to-one matching task described by Frank et al. (2008).

The ratios of correct versus incorrect responses in Frank et al. (2008) can be directly contrasted with those gleaned from ours. (Gordon [2004] does not present such ratios.) For the hidden matching task, 24/56 trials in their study contained correct responses, as did 24/56 trials in our study. For the orthogonal matching task, 24/56 trials in their study contained correct responses, as did 24/56 trials in our study. For both control tasks, then, the ratio of correct to incorrect responses was identical across the studies. For the one-to-one matching task, however, 54/56 trials in Frank et al. contained correct responses, whereas only 32/56

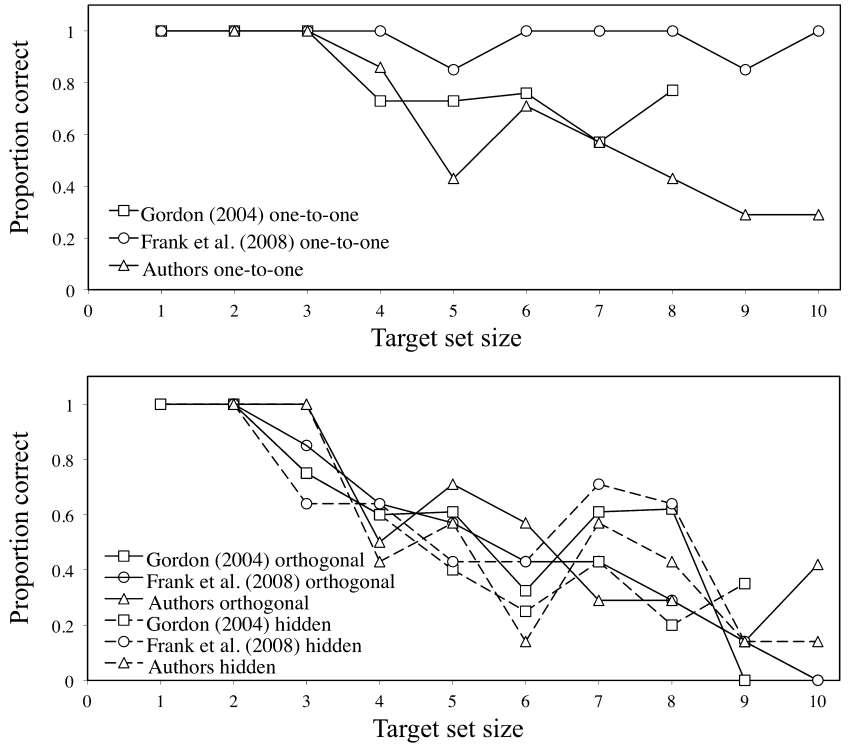

Fig. 1. Comparison of proportions of correct responses in the one-to-one line-match task and control tasks, respectively, according to number of target stimuli presented to participants.

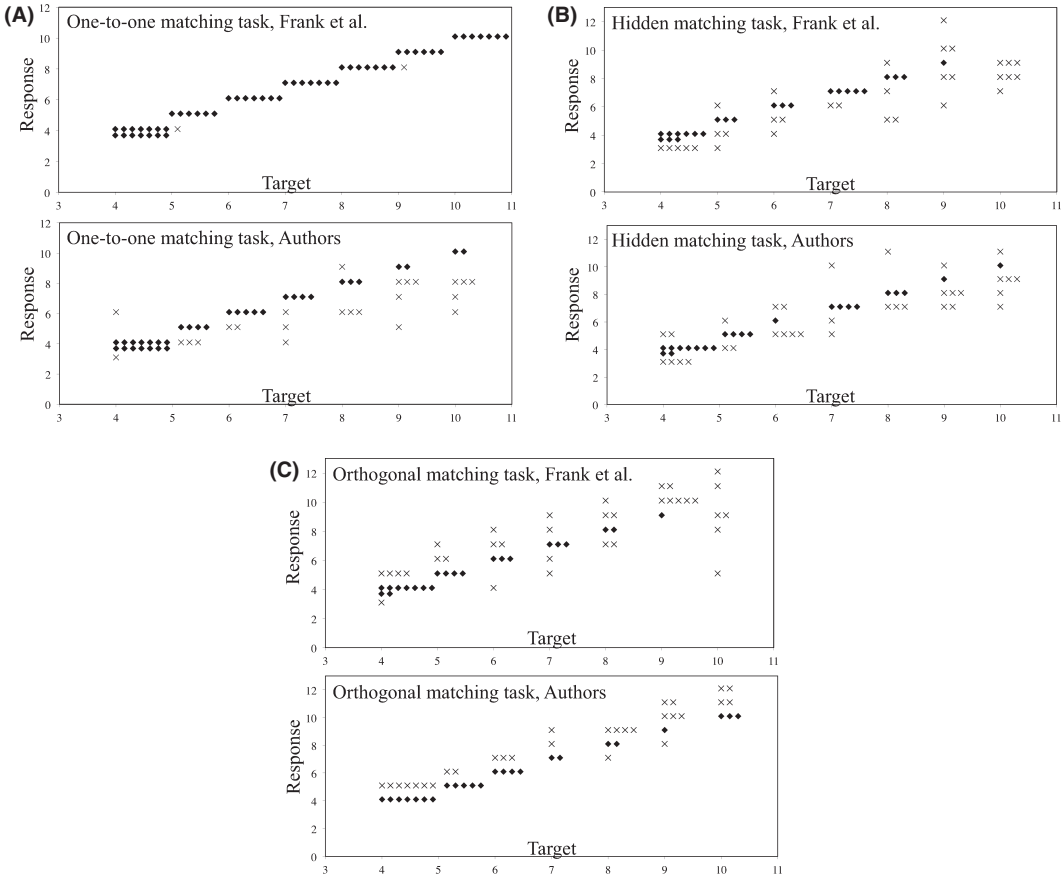

Fig. 2. Performance on individual trials for (a) one-to-one matching tasks, (b) hidden matching tasks, and (c) orthogonal matching tasks in Frank et al. (2008) and in this study. For each task, the results from this study appear below those of the corresponding results from the previous study. The *x*-axes plot the number of spools of thread presented for each trial, and the *y*-axes plot the number of balloons used by the Pirahã participants when attempting to match the target array. Correct responses are marked with a diamond; incorrect responses are marked with an *x*. Multiple correct or incorrect responses for a particular target quantity are staggered.

did in our study. Crucially, in the previous study 13 of 14 participants presented no errors for the one-to-one matching task, whereas only 2 of 14 participants in ours presented no errors.

Figure 2 contains charts depicting the performance of the Pirahã on each of the trials for the three tasks, for our data and for those in Frank et al. (2008). As can be seen in the figure, the performance on the Pirahã across both studies is broadly similar on the control tasks, with numerous errors that generally increase in magnitude as the number of stimuli increase. This characterization is also accurate for our one-to-one matching task.

In Frank et al. (2008), the performance of the Pirahã on the one-to-one task differed significantly from each control task when participants' percent correct responses were contrasted ($p < .001$ in each case). For our results, the one-to-one task and the hidden matching task did not differ significantly according to this metric, paired $t(13) = 1.65$, $p = .123$, nor did the one-to-one task and the orthogonal task, $t(13) = 1.53$, $p = .150$. Our results for the one-to-one task differed significantly from their one-to-one results, however, according to this measure, $t(13) = 6.62$, $p = .000$. The mean of all responses for our one-to-one matching task also differed significantly from that in Frank et al., $t(55) = 4.06$, $p = .000$. Differences in responses for the two control tasks were comparatively minor. For the hidden matching task means, differences between the two studies' results did not attain significance, $t(55) = 1.21$, $p = .233$. The same is true of orthogonal matching task, $t(55) = 1.92$, $p = .061$).[1] A motivation for the cross-study differences is presented in Section 4.

The coefficient of variation (CoV; standard deviation divided by mean, calculated for each set size) has been shown to hover around 0.15 when analog estimation is used in such tasks (Weber's law). Such analog estimation is characteristic of the implementation of the approximate number system, discussed in works such as Carey (2001) and Condry and Spelke (2008). The aggregate CoV across all the tasks in Gordon (2004) was 0.15. The CoV for our one-to-one matching task also averaged 0.15 for all set sizes. As can be seen in Fig. 3, similar CoV's were obtained for the two control tasks described here, both in our study and in Frank et al. (2008). Only the CoV for the one-to-one matching task in the latter study differed markedly from the other CoV's obtained to date in research among the Pirahã. Put simply, the CoV figures are generally indicative of the Pirahãs' employment of analog estimation strategies when attempting to match quantities > 3.

3. Experiment 2: Recognition of already-matched quantities

All of the number-recognition tasks so far documented among the Pirahã involve the elicitation of some presentation of a certain number of objects by the participants. One could argue that it is possible that the Pirahã may recognize exact quantities over 3, but that when asked to produce a matching array of objects, they defer to approximation strategies. Such a strategy would give the impression that the people were incapable of exact recognition, when they perhaps simply avoid it in such contexts. Conversely, it is also possible that the performative matching tasks may actually prevent the Pirahãs' from making certain kinds of errors. Consider that for the one-to-one matching task in our study and for the hidden

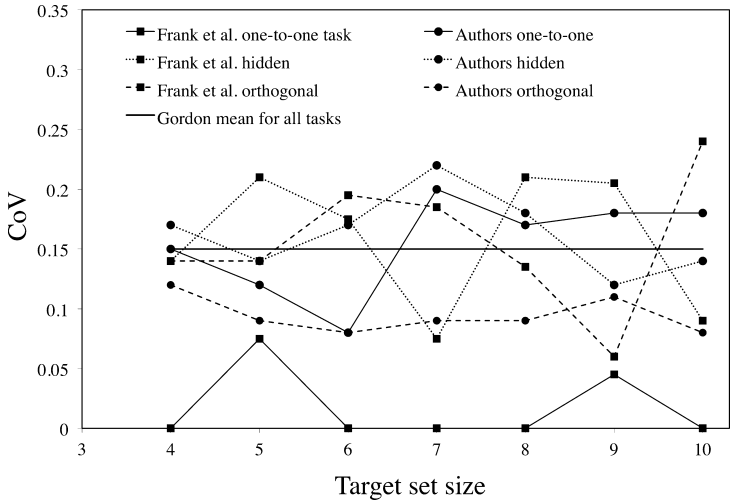

Fig. 3. Contrast of coefficients of variation across three studies.

matching tasks in our study and in Frank et al. (2008), the people rarely overestimated the number of stimuli. This is likely due to basic factors of spatial recognition; that is, the Pirahã stop presenting balloons when the line of balloons begins to exceed in length the line of spools. To account for such factors, we conducted another simple task. For the task, the same 14 Pirahã tested in the first experiment were presented with two lines of stimuli, one of balloons and one of spools of thread. The lines of stimuli were equal in length. In some cases, however, the number of balloons exceeded the number of spools (by one), in other cases the lines were equal in number, and in other cases the number of balloons was smaller than the number of spools (by one). The Pirahã were asked whether the lines were the ''same'' or ''not the same'' in amount by K.M. in Pirahã, after K.M. and C.E. first modeled the task in Pirahã, with matched and non-matched lines. The results of the trials for this task are depicted in Fig. 4. Two participants were not tested after they presented errors on the trial tasks involving sets of two and three while generally appearing uncertain about their responses in a way the other participants did not. The remaining 12 participants grasped the task, as evidenced by correct responses on such trials with lower numbers. As in the case of the first experiment, participants were tested on a subset of at least four numbers (with smaller and larger numerosities) to avoid effects from fatigue or disinterest.

The responses of all participants contained at least one error for greater numerosities. In some cases the participants judged unequal quantities to be the same, and in others they judged equal quantities to be different. There was a modest tendency for correct responses when the two lines were equal in number, although overall performance on the task was extremely poor. This poor performance cannot simply be ascribed to poor task comprehension, however, since incorrect responses in the case of all 12 participants only surfaced for numbers > 3. The results suggest that the Pirahã face extreme difficulties in recognizing exact quantities > 3 if the quantities differ by only one. They also suggest that their

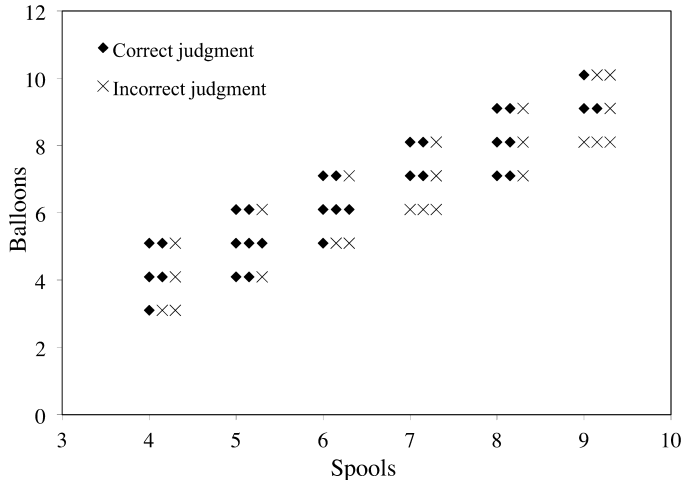

Fig. 4. Results of non-performative recognition task based on judgments of ''sameness.''

responses for the tasks in the first experiment may actually have benefited from their ability to performatively match the lines of objects.[2]

4. Discussion

Some elaboration of the cultural ecology is important here, so that we can better frame the results of all three studies. The Pirahã are a group of approximately 700 semi-nomadic people that live along the Maici River, a tertiary tributary of the Amazon. They live in small villages, generally with about 10–15 adults. Three of these villages are located near the Trans-Amazon highway (in actuality a dirt road), as is apparent in Fig. 5. Two of these villages are located about 10 and 20 km from the road and were the sites of the experiments described in Gordon (2004). The other is a new village located adjacent to the road, where this study was undertaken. A fourth village is the Pirahã-named ''Xagiopai'' village, where the experiments in Frank et al. (2008) were run.

While the Pirahã are monolingual, they have regular contact with outsiders. For the two villages near the road, this contact consists primarily of interactions with transients and a few owners of land nearby. In the Xagiopai village, however, there has been a consistent presence of Brazilians and others for the last decade. A government-run clinic has been maintained in the village until recently, and several Western structures have been built there. Significantly, in the 2 years preceding the field experiments of Frank et al. (2008), K.M. spent months in the village. During this period, she sought to teach basic arithmetic to the people at their request. K.M. also used one-to-one matching tasks in her arithmetic sessions with the Pirahã at the Xagiopai village. In order to foster improvement in the direct one-to-one matching task, K.M. innovated several words, employing Pirahã morphemes, for numbers between 4 and 10. For example, she introduced the word *xohóíśógió*, meaning ''all

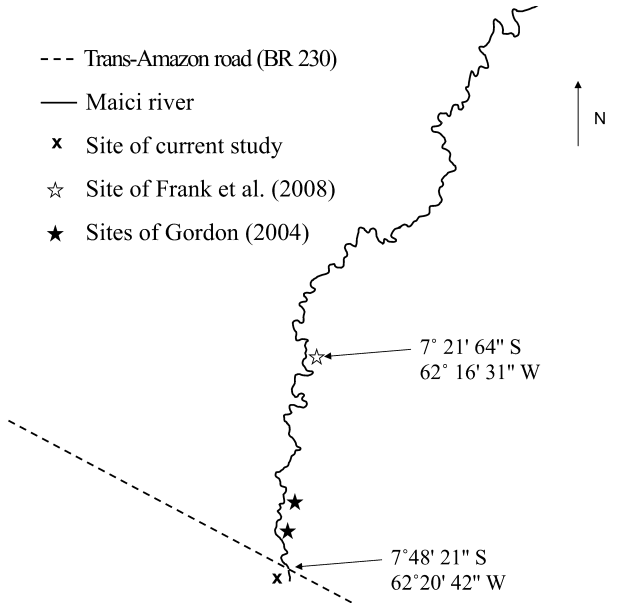

Fig. 5. Outline of the Maici River from the Trans-Amazon in the south to the mouth of the river in the north. The four locations relevant to this discussion are marked.

the sons of the hand,'' and used it to refer to the number 4. Anecdotally, she notes that she witnessed qualitative improvement in the performance of the Xagiopai Pirahã on the one-to-one matching task after introducing novel number words such as *xohóíśógió*, and not after utilizing other tactics (e.g., positive feedback). All of the adults at Xagiopai participated in such tasks numerous times in 2006, in the months leading up to the field research on which the findings in Frank et al. were based. (K.M.'s presence in the village during this period is documented via flight logs.) We are able to catalog the names of the tribe members in question and have corroborated their participation in the research conducted for Frank et al. The Pirahãs' heightened performance on the one-to-one matching task in that study can plausibly be explained because of their exposure to the task in question, along with their exposure to the innovated words. In principle, it is difficult to dissociate the potential conflating effects of their exposure to the task and the neologisms in question. It is worth noting then that Pirahãs at other villages, including those tested in Gordon (2004), had been exposed to the one-to-one matching task and other attempts at basic math training (Everett, 2005, p. 625). It was only at Xagiopai that the adults were exposed to both the task and the innovated number words, suggesting quite plausibly that their exposure to such number words, not merely their familiarity with the task, motivated their enhanced performance.

An alternate account is that Frank et al.'s (2008) results perhaps differed because the Pirahã understood the tasks better for their study; that is, there was greater clarity of instruction for their experiments. There are several issues that militate against this alternate account. The first is that, for this study, the Pirahãs' instructions' were provided by a field

worker who has documented experience speaking their language for over 30 years. The second issue is that we replicated exactly the methods described in Frank et al. Third, the relatively constant CoV evident in our results for the first experiment is not suggestive of a mere lack of understanding of the tasks by the Pirahã. It is suggestive instead of a pattern of analog estimation. Fourth, the performance of the Pirahã on the control tasks did not differ significantly across the two studies. These four issues would seem to make the alternate account untenable.

What seems quite clear is that the ''one-to-one match'' findings in Frank et al. (2008) have not been replicated here. We do not claim that the results obtained by Frank et al. vis-à-vis the one-to-one matching abilities are inaccurate, merely that they appear to reflect the abilities of Pirahã in only one of the four villages tested, one in which a history of number training, number–word usage, and general exposure to Western contact is documented. If the motivation we suggest for Frank et al.'s disparate results on the one-to-one task is correct, their findings on that task are crucial to our understanding of Pirahã numerical cognition since they are consistent with the notion that the Pirahã can exactly match quantities > 3 once they are adequately familiarized with number terms denoting such quantities.

4.1. Further evidence of atypical Xagiopai results

The Xagiopai village is currently not accessible to non-government workers. We were fortuitous in that, at the time of our research, two members of the Xagiopai village were visiting the location of the current test site. Their visit overlapped only briefly with ours; nevertheless, we were able to conduct the crucial one-to-one matching task with these two additional participants. Both of these members had previously worked with K.M. and recalled her work with innovated number words. We tested these Xagiopai residents on the crucial one-to-one matching task for all numbers from 4 to 10. Crucially, they presented no errors for any of the trials. These results further suggest that the Xagiopai Pirahã outperform the remaining Pirahã on the pivotal one-to-one matching task, while further suggesting as well that our results for the non-Xagiopai Pirahã were not due to methodological factors.

4.2. Conclusion

Based on the results here and in Gordon (2004), it appears that the Pirahã at villages besides Xagiopai, that is, those who have not had extensive training with one-to-one tasks while being exposed to innovated of number words, are *not* able to consistently recognize exact quantities over 3, even when no recall or spatial transposition is required.

K.M. did not conduct experiments of the sort documented here during the period in which she innovated number words at Xagiopai in 2006, and so we are left with anecdotal evidence only that she witnessed a qualitative improvement in the Pirahãs' performance on the one-to-one task after she introduced number words. Nevertheless, given the heightened performance of the Xagiopai Pirahã for the one-to-one matching task, and given that her account is based on extensive contact (decades) with the people, we feel that it provides us with the most plausible motivation for the discrepant performance of the Xagiopai group

documented in Frank et al. (2008). Crucially, however, our conclusion does not rest on this claim. It rests instead on the finding that the accurate performance on the one-to-one matching task has only been observed in one Pirahã village with an atypical history of outside contact. Most Pirahã face difficulty when asked to match exact quantities > 3, and this difficulty surfaces in all tasks so far conducted among the people. The most plausible motivation for this difficulty is, we believe, that they do not have access to the crucial ''conceptual tool'' (see Frank et al., 2008; Gentner, 2003; Gentner & Christie, 2012) of number terminology.[3] Judging from the results obtained among this population, this conceptual tool is essential to accurate performance on various quantity recognition tasks. It is clearly required for tasks requiring transposition or recollection of stimuli (as suggested by Frank et al., 2008), but apparently it is also required for the basic one-to-one matching task not involving additional cognitive demands.

The results in Pica, Lemer, Izard, and Dehaene (2004) suggest the speakers of Mundurukú, another Amazonian language with a modest inventory of number words, also face difficulties in exactly matching quantities over 3. The data from Mundurukú and from Pirahã speakers are generally consistent with the suggestion made in Dehaene, Spelke, Pinel, Stanescu, and Tsivkin (1999) and Lemer, Dehaene, Spelke, and Cohen (2003), inter alia, that there are two distinct neurophysiological non-verbal systems employed for large number approximation and for the recognition of small quantities, respectively. They are also generally consistent with the evidence discussed in Condry and Spelke (2008) and elsewhere that language may serve to conjoin these two non-verbal systems for the purposes of exactly recognizing and manipulating unbounded numerosities. At the most basic level, the data presented here are consistent with the notion that the exact recognition of quantities > 3 relies on a culturally constructed conceptual tool, namely precise number terminology, which is not universal to all human societies.

Notes

1. For the orthogonal matching task, the Pirahã responses were generally characterized by overestimation. In the case of the orthogonal task in Frank et al. (2008), this tendency was slightly less pronounced, especially for 10 stimuli. If trials involving 10 stimuli are not considered, the disparity between the means of the orthogonal tasks becomes even less apparent, $t(48)=1.42$, $p = .162$.
2. A reviewer notes that participants may have perceived some differences but not considered them meaningful enough to report for higher numerosities. We cannot rule out such an interpretation, but note that it would still imply the participants' utilization of approximation strategies for numerosities > 3. It is unclear why they would volitionally adopt such a strategy when the experimenters modeled a non-approximate strategy.
3. Everett (2005) suggests that they lack this conceptual tool because of more general cultural constraints.

Acknowledgments

The authors wish to thank Daniel Russell, Jamie Scotti, Daniel Everett, and most crucially the Pirahã who participated in this study. They are grateful as well for comments offered by Phillip Wolff, Peter Gordon, Michael Frank, Edward Gibson, and Evelina Fedorenko. The first author was supported by a Provost Research Award from the University of Miami.

References

Carey, S. (2001). Cognitive foundations of arithmetic: Evolution and ontogenesis. *Mind and Language*, *16*(5), 37–55.

Condry, K., & Spelke, E. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology: General*, *137*(1), 22–38.

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*(4), 437–441.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*(5416), 970–974.

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.

Frank, M., Everett, D., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*(3), 819–824.

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195–236). Cambridge, MA: MIT Press.

Gentner, D., & Christie, S. (2012). Language and cognition in development. In M. Spivey, K. McRae, & M. Joanisse (Eds.), *The Cambridge handbook of psycholinguistics*. Cambridge, England: Cambridge University Press.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*(5695), 496–499.

Hammarström, H. (2010). Rarities in numeral systems. In J. Wohlgemuth & M. Cysouw (Eds.), *Rethinking universals: How rarities affect linguistic theory* (pp. 11–60). Berlin: Mouton De Gruyter.

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia*, *41*(14), 1942–1958.

Lipton, J. S., & Spelke, E. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science*, *14*(5), 396–401.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503.

Sarnecka, B., & Gelman, S. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition*, *92*(3), 329–352.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*(6389), 749–750.

Xu, F., & Spelke, E. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11.