Animal language studies: What happened?

Irene M. Pepperberg 1,*

Phone: 617-495-2852
Email: impepper@wjh.harvard.edu
Email: impepper@media.mit.edu

1 Department of Psychology, William James Hall, Harvard University, 33 Kirkland Street, Cambridge, MA, 02138 USA

Abstract

Keywords

Animal cognition
Interspecies communication
Animal language studies
Language evolution

Preparation of this article was supported by donors to The Alex Foundation.

What follows is an opinion piece, from someone who was present and part of the field almost from its inception, who attended the relevant conferences, who experienced first-hand the interactions between the major players and the interactions among these players, the press, and the scientific community at large. As such, it is a personal view of what happened; it is not meant to be a thorough scientific review of all the experiments, criticisms, and rebuttals, or even a thorough review of where we are today. I accept and acknowledge that others will have their own memories, their own interpretations, and their own views.
The 1960s through the mid 1970s was an incredibly exciting time to be involved in studies on animal abilities. For the first time, the Nobel Prize in Physiology or Medicine (1973) had been won by three ethologists—von Frisch, Lorenz, and Tinbergen. The field of psychology had been shaken up by the aptly named “cognitive revolution”—the radical notion that levels and types of intelligence in nonhumans formed a continuum with those of humans (e.g., Hulse, Fowler, & Honig, 1968). These events inspired researchers to study a wide range of behavior—including communication—in various species. Griffin (1976) encouraged his colleagues by arguing that interspecies communication would be “a possible window on the minds of animals” (p.100-105, 154, 171). Moreover, what came to be known as “animal language studies” were also in progress. Previous attempts to teach language to apes (e.g., Hayes & Nissen, 1956/1971; Kellogg, 1968) and dolphins (Lilly, 1967), using human speech, had failed. However, researchers had hypothesized that alternative modes of communication might prove fruitful, and were achieving some success. Different labs were using different techniques to establish two-way communication with our nearest relatives—chimpanzees (R. A. Gardner & Gardner, 1969; Premack, 1971; Rumbaugh et al., 1973), an orangutan (Miles, 1978), and a gorilla (Patterson, 1978). Other labs were investigating similar abilities in nonhuman species not at all closely related to humans (e.g., Herman’s work on dolphins). And the amazing studies on vocal learning in songbirds, showing the striking comparisons between the process of song acquisition and human language learning (e.g., Marler, 1973), inspired me to begin work on training a Grey parrot to use referential speech (Pepperberg, 1981).

Results from the different laboratories were divergent, but complementary. Use of American Sign Language and Signed English (Gardners, Miles, Patterson) allowed for flexibility, innovation, and direct comparisons of communicative acquisition between child and ape. Use of plastic chips to represent labels, taught via a no-fault choice procedure (Premack), provided less information about communication skills but began to elucidate how acquisition of symbolic representation could affect cognitive processing. Use of a glorified Skinner box, initially with an ape named Lana (Rumbaugh), removed most of the effects of social interaction to get at which basic concepts could be acquired via associative learning and how such learning could still allow for innovation. Herman’s early work ran afoul of animal rights activists, but with new subjects he had begun to...
show that dolphins could respond to specific cues with specific actions that demonstrated referential comprehension. My parrots started to use the sounds of English speech to identify objects, colors, and shapes. A media storm ensued (e.g., NOVA and BBC Horizon did documentaries; numerous articles were published in places like The New York Times). Not only had we achieved a kind of “Dr. Doolittle” moment, but we felt we could be gaining insights into how language and complex cognition might have evolved in our ancestors. If creatures separated by 300 million years of evolution and with remarkably different-looking brains could all acquire some level of symbolic representation and regular ordering of those symbols, wouldn’t that imply something basic in evolution? How might our ancestors have built upon such abilities?

These studies were even more exciting because they began only a decade or so after major competing theories had been proposed for how children acquired language—Chomsky’s (1959) innate Language Acquisition Device (LAD) and Skinner’s (1957) tabula rasa, in which conditioning played a major role. Child developmental laboratories sprung up, gathering data to try to support one side or the other (amazingly, little research had actually previously been done), but the implications of the studies on nonhumans were clear: To state the case in the simplest terms possible, if Chomsky was correct, no nonhuman could possibly acquire anything like language; if Skinner was correct, it was only a matter of time, energy, effort, and the correct procedure that would ensure success.

The stakes were high, and many researchers wrote scholarly articles questioning what exactly nonhumans had learned, disputing the extent of the claims being made (e.g., Bronowski & Bellugi, 1970; Lachman & Mister-Lachman, 1974; Lenneberg, cited in Nottebohm, 1973). These articles, and their rebuttals, started serious discussions of truly important questions: for example, what were the actual hallmarks of language; what might the apes’, dolphins’, and parrot’s abilities tell us about language evolution and cognitive processing; what stages did children go through en route to full language; how did codes such as ASL differ from spoken language, and were these differences important? (Note that at one point some scientists questioned if ASL was even a real human language; a full analysis hadn’t been published until Stokoe, 1978.) If nothing else, data from these studies spurred research on child language acquisition and cognitive development. As more was learned, the bar kept being raised for the nonhumans: Once nonhumans could use symbols to refer to objects, they needed to use
symbols for verbs, then needed to construct phrases, and also needed to use this acquired code to demonstrate complex cognitive processes (categorization, relational concepts, same-different, etc.). In frustration, Fouts (1974) basically argued that language seemed to be defined as whatever it was that apes didn’t have. Nevertheless, what our animals did learn provided important insights. Whether or not Premack’s plastic chip system could be called “language,” only apes who had undergone such training seemingly could learn concepts such as formal analogies (Premack, 1976). Did such training actually alter the apes’ brains? Could these data provide information on how symbolic representation, cognitive processing, and brain development might have interacted to make changes in our ancestors en route to modern humans?

At this point, however, no one had argued that problems existed with the data being collected. At least not until Terrace (Terrace, 1979a; Terrace et al., 1979). Terrace reported that his ape, Nim (named Nim Chimpsey, in a stab at Chomsky), learned very little after being trained in sign language but via the techniques of operant conditioning (Fouts, 1983). Other signing apes had been taught in ways based on those used with young children: rich in social interaction, modeling, and referential rewards (if an ape signed something about X, it usually received X or got to do X; e.g., R. A. Gardner & Gardner, 1969). Terrace and colleagues (1979) argued that because his ape could not create a sentence comparable to one used by adult, oral humans, no ape could acquire anything vaguely resembling human language. Although Terrace (1979a) did raise points about methodology and data interpretation that needed addressing (e.g., the sometimes small numbers of options from which subjects could choose the correct answer and often the small numbers of trials involved, both lowering the statistical power), he did not limit his criticisms to those points. Specifically, he compared ASL-learning chimps not with ASL-learning human infants, but (improperly) with infants acquiring spoken English. Thus, he correctly noted that ASL strings such as YOU ME EAT? lacked the complexity of English syntax, but failed to acknowledge that such strings were one way in which human oral sentences such as “Wanna grab lunch?” might be expressed in sign, particularly by children, and thus were, at many levels, equivalent (note B. T. Gardner & Gardner, 1998; Van Cantfort & Rimpau, 1982). He didn’t accept that in a system in which one can’t increase volume, emphasis occurs by perseveration (Finton & Smith, 2003; B. T. Gardner & Gardner, 1998; Hoffmeister, Moores, & Ellenberger, 1975). He acknowledged that his ape never progressed beyond simple associations between
objects, a few actions, and symbols, but not that such associations are often the very first stages in human label acquisition (e.g., Bloom, 1973). (Numerous reasons likely existed for Nim’s failures, some of which might have been the huge number, ~40, of different trainers, few of whom were proficient in sign; or the use of nonreferential food rewards for the acquisition of nonfood signs, thus making the association between sign and object less relevant; see Terrace, 1979b). He did not accept evidence presented by other laboratories for spontaneous generation of signs (“cry hurt food” for radishes; Fouts, 1974), or concatenation of computer lexigrams (“coke that is orange” for Fanta; Rumbaugh, 1977). Terrace went even further in his 1979 articles, however, arguing that the other signing studies were no better than his, that their data were not being analyzed appropriately, and that such studies were essentially worthless.

What could have been a series of academic arguments (see, e.g., the exchanges between Schusterman and Herman in the late 1980s, cited below), even resulting in collaborative efforts (e.g., combining the strengths and eliminating the weaknesses of the varying training and testing techniques to determine something about language primitives and nonhuman cognitive processing), instead devolved into total chaos when the New York Academy of Sciences, in 1980, hosted a conference put together by Thomas Sebeok, a noted researcher in zoosematics, and Robert Rosenthal, who studied nonverbal communication and how expectancies influenced conclusions (Sebeok & Rosenthal, 1981). The conference not only had scientists as speakers but also nonscientists like the “Amazing Randi,” a professional magician, who demonstrated how easily people could be fooled into seeing what they wanted or expected to see. Sign language researchers were accused of cuing their apes by ostensive signals (even though apes, as it turns out, may have some difficulty interpreting such forms of human action; see Bräuer, Kaminski, Riedel, Call, & Tomasello, 2006), and of consistently overinterpreting the animals’ signs (a possibility in some instances, but not in others; see R. A. Gardner & Gardner, 1984, for controlled vocabulary tests). Scientists such as the Rumbaughs vehemently objected to the assault, but at the time argued that only their own computer-based system prevented the problems the conference was addressing, thus showing little affiliation with, or giving any support to, researchers using other techniques (Marx, 1980; Wade, 1980). They thus added to the furor sparked by Sebeok and Rosenthal, who all but called researchers in the field liars, cheats, and frauds (and actually did so in
a postconference press gathering; Wade, 1980). The conference was covered by
the media (see review for *Science*; Wade, 1980), and the public brouhaha meant that government agencies—responsive to
the blow-back—fairly quickly cut off the funding for all of the studies.

Almost all of the laboratories abandoned their studies of language *per se*, but the
silver lining was that most shifted to using what we called “two-way
communication systems” to examine various forms of cognitive processing that
relied on symbolic representation—for example, studies on numerical concepts
(e.g., Boysen, 1993; Boysen et al., 1993; Matsuzawa, 2009; Pepperberg, 2006;
Pepperberg & Carey, 2012); rule-governed behavior, perception, and cognition
(e.g., Herman, 1987, 1988, 2010; Herman et al., 1993; Schusterman & Gisiner,
1988, 1989; Schusterman & Krieger, 1984); relational concepts (Pepperberg &
Brezinsky, 1991; Schusterman & Krieger, 1986); and symbolic equivalence
(e.g., Kastak & Schusterman, 2002; Pepperberg & Gordon, 2005; Reiss &
McCowan, 1993). Other researchers examined different forms of rule-governed
behavior (and often social learning) that were the basis for syntax, such as the
ordered sets of actions needed to solve puzzle boxes and how those ordered sets
may be acquired (keas: Miyata, Gajdon, Huber, & Fujita, 2011; apes: Whiten,
1998; note Terrace, Son, & Brannon, 2003, on other forms of serial learning in
monkeys).

The Rumbaugh, at least for awhile, continued to examine aspects of symbolic
communication, although primarily in terms of comparative cognition, studying
similarities and differences between bonobos and common chimpanzees with
respect to aspects of symbolic labeling and comprehension of sentence frames
(reviewed in Savage-Rumbaugh, Brakke, & Hutchins, 1992). Interestingly, this
later work often specifically tested effects of social interaction, acknowledging a
significant change in methodology and interpretation. Other researchers
examined receptive capacities of dogs, possibly to see if the process of
domestication had an effect on referential learning (e.g., Kaminski, Call, &
Fischer, 2004; note Griebel & Oller, 2012) and still others worked to “crack the
code” of communication in nature (e.g., monkeys: Schlenker, Chemla, Arnold, &
Zuberbühler, 2016; song sparrows: Beecher & Akçay, 2014; toothed whales:
Janik, Sayigh, & Wells, 2006; McCowan & Reiss, 2001). Some researchers have
looked for rule-governed behavior with respect to communication systems by
studying nonhumans’ (particularly songbirds’) understanding of artificial grammars (Beckers, Bolhuis, Okanoya, & Bewick, 2012; Gentner, Fenn, Margoliash, & Nussbaum, 2006; Fitch & Friederici 2012; ten Cate & Okanoya, 2012).

Thus, although interspecies communication studies may not have taught nonhumans to use “language,” nonhumans (and those who studied them) had learned quite a bit. Clearly, some common neural architecture enabled disparate nonhuman species to achieve a level of symbolic representation and rule-governed behavior, suggesting that some such abilities were likely in their natural communication systems and had evolved for that purpose—it was unlikely that researchers instilled such behavior entirely de novo. Nevertheless, grants for such studies became more and more difficult to obtain, subjects began to die off, and the next generation of students seemed more intrigued by human neurobiology and fMRIs. And, in a somewhat unnerving turn of events, animal rights leaders began using the data obtained by researchers in the animal language/animal cognition field to push for rules and regulations prohibiting apes and marine mammals from being available for such studies in the future.

As a consequence, we are missing many opportunities. The possibilities of studying nonhumans as models for our ancestral abilities or for how convergent evolution might have led to similar language-like abilities in birds, humans, and marine mammals has significantly decreased—as is the possibility of beginning studies with other vocal learners such as elephants (note Stoeger et al., 2012). Fortunately, the field of animal cognition still has pockets of strength, and questions abound: What are the differences in imitative ability (in all its complexity—e.g., mimicry vs. emulation vs. imitation) that might be important for human–nonhuman communication systems and acquisition of cognitive concepts (e.g., Nielsen, Subiaul, Galef, Zentall, & Whiten, 2012)? Given that we know more now about signed languages, their parallels to spoken languages, and how they can evolve over time (e.g., Goldin-Meadow et al., 2015), and more about language pedagogy (e.g., Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015), how much might apes learn if trained appropriately? Given the research on using computer-based communication training and portable systems for children on the autistic spectrum (Ramdoss et al., 2011), where might such training have led with apes and marine mammals? Given that we know more about human languages and are continuing to learn more (e.g., Levinson & Gray,
are the criteria we once used for nonhuman acquisition fair? With data on those fronts, not to mention knowledge of brain structures and connectivity, and the striking parallels between primate and nonprimate and even nonmammalian communication systems, what might we be able to deduce about the precursors to modern human languages? Might studies of avian vocal learning—with respect to possible avian “missing links,” such as bellbirds (flycatchers technically classified as suboscines—nonvocal learners—that nevertheless do learn simple songs; Kroodsma et al., 2013)—tell us, through investigations of convergent evolution, something about what types of brains our ancestors might have been developing? The list can go on, and the implications are clear.

References


(Baby)Talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science, 24*, 339–344.


Savage-Rumbaugh, E. S., Brakke, K., & Hutchins, S. (1992). Linguistic development: Contrasts between co-reared *Pan troglodytes* and *Pan paniscus*. 
In T. Nishida, W. C. McGrew, P. Marler, M. Pickford, & F. B. M. de Waal (Eds.), *Topics in primatology: Human origins* (pp. 51–66). Tokyo, Japan: University of Tokyo Press.


