



Metasignaling and Language Origins

ABSTRACT Over the past two decades, comparative linguistic anthropological research has disclosed the significance of meta-discourse and metasignaling for the establishment, maintenance, and transformation of social relations. One type of metasignal–signal relationship has, because of its role in producing sociability, come to seem especially interesting with regard to the evolutionary origins of human language—the type wherein a strategically (and presumably neocortically) induced metasignal is modeled after, but distinct from, an instinctively triggered signal, as in ritualized lamentation. Strategic vocal deceptions in nonhuman primates are possible precursors of true socially constructed, socially shared metasignals, which in turn may be ancestors of modern human language. This article charts an evolutionary path from strategic vocal manipulation, to cultural metasignals, to language, using data from primate studies, discourse-centered research, and early child language acquisition. [Key words: language origins, semiotics, metacommunication, chimpanzees, early language acquisition]

FOSSIL POETRY—that was how Emerson described language—the residuum of potent communications preserved in outline form, lacking in lifeblood. The search for language origins, from this perspective, is the search for the origins of fossils, rather than for the once living communicative phenomena that produced them. My goal here is to resuscitate the living phenomena behind the fossils—the phenomena that would be ethnographically describable communicative social interactions—and to draw on developments of the last few decades in the area of discourse-centered approaches to culture to provide a framework in which to think about the origins of language. My specific thesis is that language originates from special communicative interactions that involve metasignaling, that is, from signals having as (at least) part of their meaning other signals.

As an example of metasignaling, consider the stylized or ritualized forms of lamentation found in many cultures around the world.¹ Such laments involve the use of the vocal apparatus to produce sounds that remind the listener of crying. And these laments are deployed as part of strategic communicative interactions and are presumably neocortically induced, rather than innately controlled.² As I have argued elsewhere (Urban 1988, 1991), however, ritual laments do not only remind the listener of crying. They are designed to appear as distinct from crying, for example, by virtue of their regular intonational contours and line lengths, which give them a musical quality. This distinctiveness from that which they resemble results in their

possession of a new social meaning. They are expressions of the desire for sociability.³

Why regard them as metasignals, rather than simply as separate signals whose meaning is the desire for sociability? My contention is that the meaning possessed by stylized laments as socially communicative signals cannot be fully disentangled from crying as a signal. What allows a listener to grasp their purpose is, in part, the way in which they suggest crying. By resembling crying, but also by being distinct from it, they guide the listener in how to reason about their meaning as social signals. If the metasignals (the ritualized lamentations) resemble the signals (crying) in their perceptible shape but are not identical to them, then, or so the listener reasons, the meaning of the lament must resemble the meaning of the cry (distress, pain, need of help) but be distinct from it (need for social connectedness). I have attempted to diagram this situation in Figure 1.

I propose to argue that primitive, language-like vocal forms emerged with the birth of such metasignals,⁴ produced and interpreted by means of neocortical reasoning. Metasignals serve as semiotic levers for the production of new signals, prying them apart from old ones. In this scenario, the first metasignals were those that, like the stylized lamentations, involved the relationship between a neocortically shaped signal and a limbically driven one. As signals proliferate, the new separate out from the old in accord with a constant principle, relying for their communicative success in part on their similarities to old signals and in part on their differences from them. The result of

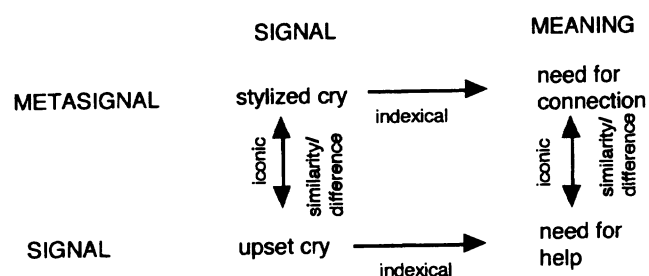


FIGURE 1. Stylized crying as metasingal. The stylized cry is similar to but different from the upset cry. This calls attention to the similarity coupled with difference between the respective meanings of the metasingal and signal, allowing individuals to reason from the stylized cry to its meaning (need for social connection) via the intermediary of the upset cry.

iteratively applying this principle is a network of interconnections among signals. That network is reminiscent of grammar, with its morphology and syntax.

DO NONHUMAN PRIMATES POSSESS THE ABILITY TO NEOCORTICALLY FASHION SIGNALS?

It is obvious that humans possess the capacity to strategically construct, deploy, and interpret signals for social purposes. Any newly uttered sentence that serves some imagined communicative end illustrates this point. But what about our closest nonhuman relatives? I submit that the evidence here too is overwhelming, especially that coming from the various attempts to teach language to chimpanzees, bonobos, or gorillas.⁵ Unarguably, our closest nonhuman primate relatives, in interacting with humans, regularly use learned signs for social ends.

What about chimps in the wild? An example of a chimp spontaneously inventing a communicative signal, apparently through strategic reasoning, is that of Mike, whose use of kerosene-cans has been documented by Jane Goodall (1968, 1986). In their attempt to assert dominance, male chimps charge their competitors, "dragging, waving, or hurling branches" and "rolling or throwing rocks" (Goodall 1986:549). Mike was smaller than the other dominant males but discovered that empty kerosene cans, rolled in front of him, could intimidate his larger rivals. By means of this innovative signal, he was able to become alpha male, despite his stature. Goodall notes that "Mike's deliberate planning was a striking aspect of his rise to alpha status":

He picked up two empty cans and, carrying them by their handles, one in each hand, walked (upright) back to his previous place, sat, and stared at the other males, who at that time were all higher ranking than himself. They were still grooming quietly and had paid no attention to him. After a moment, Mike began to rock almost imperceptibly from side to side, his hair very slightly erect. The other males continued to ignore him. Gradually Mike rocked more vigorously, his hair became fully erect, and uttering pant-hoots he suddenly charged directly toward his superiors, hitting the cans ahead of him. The other males fled. Sometimes Mike repeated this performance as many as four times in succession, waiting until his rivals had

started to groom once more before again charging toward them. When he eventually stopped . . . they sometimes returned and with submissive gestures began to groom Mike. [1986:426]

Mike seems to have reasoned his way to the construction of this signal, which fit in iconically with the stone rolling more typically employed but which had the added advantage of loud, clattering, and apparently frightening noise. Goodall notes that "while all of Mike's contemporaries had the opportunity to use these cans, and most of them occasionally did so, only Mike profited from the experience and used it to further his own ends" (1986:426).

Other examples of strategically reasoned signal construction and deployment could be given, but this perhaps suffices to demonstrate that the role of the neocortex in fashioning communicative signals is by no means unknown in the chimp world. Mike had an objective—the intimidation of other more dominant males along the way to his achievement of dominant status—that was seriously limited by his physical size and strength. However, through apparent ingenuity he was able to construct signals that allowed him to overcome his physical shortcoming. I should remark that there is no evidence in this example of the other chimps using their own intelligence to decode the signal. They seem to have responded instinctively, as if operating under limbic control. From the perspective of ethnographically describable communicative interactions in humans—including ritual lamentation—this is a key point, and I will return to it subsequently.

DO NONHUMAN PRIMATES POSSESS THE ABILITY TO NEOCORTICALLY FASHION VOCAL SIGNALS?

It is one thing to ask whether nonhuman primates have the ability to use reasoning in the deployment of signals generally; it is quite another to ask just how much control they have over vocal signals, in particular. I will not beat around the bush with my answer to the latter question. Neocortical control over the vocal tract and, in particular, the larynx is, in my assessment, the central feature that made language possible. This is because the larynx in chimps and other nonhuman primates seems, in large measure, innately programmed. It is under limbic control, going through its sequences pretty much automatically once it is triggered. However, evidence also suggests some neocortical involvement, and this seems to me the window of opportunity in making the transition to the human line and, eventually, to language. It would not be inconsistent with the view expressed here, as will become clear later, that some measure of neocortical control over the larynx may date back to the earliest differentiation of the human line from other primates, that is, back to Australopithecines, and that it may have older roots.

Keith and Cathy Hayes's (Hayes 1951; Hayes and Hayes 1952) work from the late 1940s with a chimp named Vicki pinpoints with some precision the specific problem chimps have in achieving neocortical control over the vocal apparatus. Cathy Hayes observes that "at

the age of thirty months [Vicki] used only three words and these not always appropriately" (1951:144). However, the film footage of Vicki reveals a considerable ability to control parts of the vocal apparatus in response to cues from Keith Hayes. She is able to produce somewhat recognizable versions of the English words *papa*, *mama*, and *cup*. She learns to control the opening and closing of her mouth to produce a [p]-like sound—although the sound appears actually to be a bilabial click or kiss-like sound, but it is a close enough imitation and she can produce it on cue. She produces as well a [t]-like sound—probably an alveolar click. She produces a velar fricative sound that is able to pass for a [k] in her articulation of *cup*. She even produces, with more difficulty, an [m]-like sound in which she has to achieve control over not only her mouth but also, in some measure, her velum. So she seems to have at least a rudimentary ability to manipulate her upper vocal tract in her efforts to imitate the humans around her and to produce language-like consonantal sounds.

Where she falls down is in her control over the larynx. She cannot seem to turn the vocal cords on and off at will. In the case of *mama*, the consonants and vowels she produces are entirely voiceless. She does not activate her vocal cords at all in the course of this articulation. The result is a phonetically highly marked form, with a low degree of perceptual salience. Vicki's pronunciation of *papa* also suffers in this regard. While her [p]s (or clicks) are voiceless, her [a]s, contra the principle of maximal salience, are also voiceless. For this reason, her pronunciation of *papa* appears as barely intelligible, perhaps only recognizable because we know beforehand what she is supposed to be saying. The same problem emerges again in the case of *cup*. The entire word is articulated without the benefit of vocal cord vibration. In addition, of course, the [k] she produces is really a voiceless velar fricative. The resulting syllable can be discerned only with some difficulty.

Indeed, there is no evidence in the Hayeses' experiments that Vicki was ever able to activate her vocal cords to imitate the usual, basic, or unmarked form of the human vowel—that is, the voiced form. They remark that, while Vicki "developed a type of play which is superficially similar to babbling . . . the sounds are produced entirely by mouth vibration without use of the larynx" (1952:108). She was able to control her upper vocal tract articulatory organs reasonably well: her lips, mouth, tongue, and even velum. She may even have been able to shape her mouth to produce different kinds of vowels. This is not clear. But she could not activate her larynx in imitation of the humans around her.

IS THERE NEUROANATOMICAL EVIDENCE FOR SEPARATE NEOCORTICAL AND LIMBIC PATHWAYS OF CONTROL OVER THE LARYNX?

I have said that metasignaling involves the relationship between a neocortically shaped signal and a limbically driven one. The idea of two differently controlled signals—

the one fashioned after the other by processes of reasoning that later become routinized—is not just an abstract construct designed to make sense of such phenomena as ritual lamentation. It is a construct that has a counterpart in neuroanatomy. There are, in humans, two separate neural pathways for laryngeal activation, as Uwe Jürgens (2001) has proposed based on studies of brain lesions as well as on experimental data on nonhuman primates.

In humans, production of true linguistic sounds originates in the motor area of the neocortex,⁶ the evolutionarily more recent outer layer of the brain that has mushroomed in the course of human evolution. Neural fibers from that area project downward through the midbrain in what is known as the pyramidal tract. Some of the motor neurons eventually connect with the nucleus ambiguus in the medulla, which in turn directly controls the laryngeal muscles. This is what allows humans to sing a musical note and to consciously manipulate its pitch.

If the facial motor area of the neocortex is destroyed, humans lose the ability to voluntarily produce linguistic sounds, though they are still capable of producing groans, whimpers, and laughs (Jürgens 2001). Destruction of the analogous brain area in the squirrel monkey produces no observable effect on vocalizations, leading to the conclusion that the neocortex is not centrally involved in the production of those vocalizations.

In nonhuman primates, call production is thought to be triggered by the anterior portion of the cingulate cortex. The cingulate cortex forms part of the evolutionarily ancient inner brain area. It is one of the structures making up the so-called limbic system and one, in particular, associated with emotion in humans. In squirrel monkeys, electrical stimulation of the anterior portion of the cingulate cortex elicits species-specific calls (Jürgens and Zwirner 1996). The cingulate cortex, rather than descending to the lower brain via the pyramidal tract, as in the case of the neocortical signal, descends instead into the central (or periaqueductal) gray of the midbrain. Jürgens and Zwirner (1996) have shown that if the central gray is chemically blocked, calls cannot be elicited by stimulation of the anterior cingulate cortex. This pathway is thus distinct from the neocortical pathway associated with linguistic sound.

Precisely what the interaction is between these pathways in ritual lamentation remains a mystery. However, it is likely that both are involved in at least some instances, for ritual laments are capable of producing or being triggered by feelings of sadness (presumably associated with the anterior cingulate cortex) while simultaneously containing partial or full linguistic sounds. The range of possible combinations of emotional expression with imitations of emotional expression, as well as with actual linguistic sounds, suggests a corresponding range of possible kinds of interaction between neocortical and limbic control over the larynx.

The question at hand is whether (or perhaps to what extent) primates other than humans have neocortical control

over the larynx. Observations in the preceding section suggest that control over the larynx in chimps is not highly developed, even though they appear to have control over their tongues and jaws. In a parallel vein, Jürgens (2001) notes that nonhuman primates have extensive neural connections between the neocortex, on the one side, and the jaw (through the trigeminal motor nucleus) and tongue (through the hypoglossus), on the other. What they lack are direct connections between the neocortex and the larynx.

At the same time, there appear to be at least some connections between the neocortex and larynx even in squirrel monkeys. Jürgens and Zwirner (1996) have studied the effects of pharmacological blocking of the squirrel monkey midbrain. Once the central (or periaqueductal) gray was blocked, the researchers were unable to stimulate the vocal cords through the cingulate pathway, but they were able to do so via the facial motor cortex.⁷ Their conclusion is that "these results point to the existence of two separate vocal fold control pathways at midbrain level: one limbic, responsible for non-verbal emotional vocal utterances, and one neocortical, responsible for the production of learned vocal patterns" (1996:2921). This suggests that the neurological connections for the kind of metasignaling proposed here were present, in at least rudimentary form, in our nonhuman primate ancestors.

WHY IS NEOCORTICAL CONTROL OVER THE LARYNX IMPORTANT ANYWAY?

Laryngeal control is a key to the origins of language because it is central to the production of syllables. If distinct syllable types can be readily produced and perceived, especially in the course of a single utterance stretch, then the conditions are established for those syllables to enter into meaningful contrast with one another, such as occurs in the classical Saussurean system. Figure 2 shows the contrast between a Saussurean model of the sign, characteristic of language, and a model of signaling characteristic of the broader animal world, based on indexical or iconic connections, where the signal is linked directly to the world by contiguity or physical similarity between it and the thing signaled. To have a contrastive system, in the

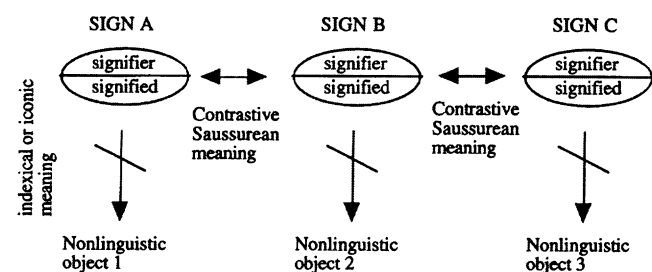


FIGURE 2. Saussurean contrastive meanings (characteristic of human language) as distinct from indexical and iconic meanings (found in animals more generally).

case of spoken language, one has to have syllables that contrast, and this requires fine-grained neocortical control over the larynx.⁸

What allows syllables to be distinguished from one another, so as to produce proto-Saussurean contrasts, is a contrast internal to the syllable itself—namely, the contrast between consonants and vowels. In a celebrated essay titled "Why 'Mama' and 'Papa'?" Jakobson (1971) observes that the words *mama* and *papa* are among the earliest recognizable linguistic productions of a child. Why should that be so? Jakobson contends that the key is the maximal perceptual difference between the consonants and vowels in the constituent syllables of these words. The maximally salient distinctions are the ones that tend to appear earliest in child development, typically by one year of age. (Note that Vicki succeeded in making utterances that resemble these two elementary words.) The distinctions are thus precursors, ontogenetically, to true language. Are they also precursors phylogenetically?

Looking at the case of *papa* a little more closely, the contrast is between a voiceless oral consonant [p] and a voiced oral vowel [a]. Indeed, [a] is the unmarked form of the vowel, found universally in all languages (Jakobson and Waugh 1987). An interesting fact about this contrast, from the point of view of stylized crying and language origins, is that it involves control over the larynx—both excitatory and inhibitory control. Human children learn to turn the vocal cords on and shut them off, even though the entire syllable may last for only one-quarter of a second (see Figure 3). The result is a syllable [pa] that is maximally recognizable as a particular type of syllable as opposed to others. Hence, laryngeal control is essential to the development of one of the most basic and universal of syllable types.

In the case of *mama*, the syllable internal contrast does not involve voicing versus voicelessness. Both the [m] sound and the [a] sound require vibration of the vocal flaps. The contrast between them, rather, is that between nasal consonant and oral vowel, the air coming out the nose in the one case and out the mouth in the other. However, the perceptual recognizability of the syllable is dependent on control over the larynx. It is crucial that both of these sounds are made with the vocal cords vibrating. One has to be able to turn the vocal cords on at the appropriate moment and then to turn them off again. If humans were not able to control the larynx—to turn it on in the first place—they would produce a voiceless nasal consonant and a voiceless vowel. Both of these phonetic types (voiceless nasal consonant and voiceless vowel) are highly marked. They occur as phonemic in few languages around the planet. They tend to be acquired, in child development, after other more basic sounds. This suggests that a voiceless version of *mama* would be less perceptually distinguishable and, hence, less likely to aid in the development of a system of Saussurean contrasts among the signals themselves. In fact, we are all aware, at some level, that this is

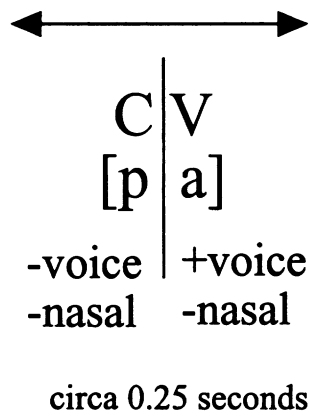


FIGURE 3. Sequential voicing contrast in the elementary syllable, where exquisite neocortical control over the vocal flaps is presupposed.

the case, for whispered words are whispered precisely so as to conceal them from others.

The conclusion I draw from this is that control over the vocal cords is essential to the development of language as a distinctive kind of communicative system. Individuals must be able to learn to turn the vocal cords on and turn them off again with split second precision. This is something that Vicki, seemingly, could not learn. If a typical syllable lasts one-quarter or one-third of a second, the production of [ma] requires that the vocal cords be turned on and then off again within that time frame. In the case of the [pa] syllable, the voiceless consonants may last only one-tenth of a second, after which the vocal cords must be made to vibrate. So the kind of control that is presupposed by stylized crying—fine-grained control over the larynx—is also presupposed by the proliferation of syllable types that is essential to the production of Saussurean contrasts and, hence, to the peculiar type of meaning associated with language. If we can understand how and why neocortical control over the larynx originated, we can begin to understand how (and perhaps why) rudimentary language, in the form of a system of syllables in contrast with one another, itself might have originated.

IS THERE EVIDENCE FOR NEOCORTICAL CONTROL OVER THE LARYNX IN CHIMPS, AND, IF SO, HOW MIGHT THAT CONTROL EVOLVE IN THE DIRECTION OF LANGUAGE?

The answer to the question, “Is there evidence for neocortical control over the larynx in chimps?” is yes. Jane Goodall describes an intriguing situation in this regard. She writes:

On one occasion when Figan was an adolescent, he waited in camp until the senior males had left and we were able to give him some bananas (he had had none before). His excited food calls quickly brought the big males racing back and Figan lost his fruit. A few days later he waited behind again, and once more received his bananas. He made no loud sounds, but the calls could be heard deep in his throat, almost causing him to gag. [1986:125]

The larynx with its vocal cords, and apparently other components of the lower vocal tract airstream mechanism, were receiving excitatory signals from the limbic or other brain areas at the same time as they were receiving inhibitory signals presumably from the neocortex.⁹ Neocortical control over vocal production, in this one dramatic instance, at least, seems to conflict with innate affective control, with strategic reasoning only partially winning out. The struggle to control innate vocalization is of evolutionary significance. In the human line, the decisive victory went to the neocortex. How might this have developed?

To answer that question, we need to know whether it is possible not only to suppress an innately triggered laryngeal pattern but also to neocortically induce the pattern. Again, Goodall supplies intriguing ethnographic evidence. Wondering whether chimpanzee “infants may learn to make use of their mothers’ protective responses” (1986:582), she writes:

My first observation of this behavior occurred as I followed Fifi and her four-year-old son, Frodo, who was being weaned. After he had twice tried to climb onto his mother’s back and twice been rejected, he followed slowly with soft hoo-whimpers. Suddenly he stopped, stared at the side of the trail, and uttered loud and urgent-sounding screams, as though suddenly terrified. Fifi, galvanized into instant action, rushed back and with a wide grin of fear gathered up her child and set off—carrying him. I was unable to see what had caused his fear response. Three days later, as I followed the same mother–infant pair, the entire sequence was repeated. And, a year later, I saw the same behavior in a different infant, Kristal, who was also being weaned.

Were these infants lying? Or was their fear real; were they suddenly frightened of maternal rejection? Obviously more observations are necessary, but I am of the opinion that they were intentionally manipulating their mothers. [1986:582]

Rather than strategically inhibiting an innately triggered cry, as in the case of Figan and the bananas, Frodo neocortically triggered a cry that also has an innate trigger. He did not modulate the cry once it had started; the cry progressed along apparently innately prescribed pathways. But he was able, evidently, to initiate the cry in the absence of the appropriate external stimulus; and he was able to manipulate it to achieve a social end. Together, the cases provide possible evidence for both neocortical excitation and neocortical inhibition of the larynx. However, the laryngeal control the chimps possess is anything but fine grained. The cries can be triggered or inhibited but not finely modulated.¹⁰

If cries are manipulated by chimpanzees to produce social connection, what would have to happen to send the chimpanzees down the road toward language? My argument is that the cries would have to become more stylized, more obviously differentiated from innately triggered cries. They would have to become like what ritual lament is today—a culturally specific form that signals the desire for sociability by resembling crying but also by resembling the specific form of crying assumed by those around one,

especially by one's parents. But how could they have become more stylized?

There is no evidence, in the case Goodall describes, that the mother responded to the cry as stylized. Instead, she reacted to it as a signal of danger. She acted accordingly to protect her child. It is true that protection consists in the mother paying attention to the child, taking it up, and carrying it off. Hence, as in ritual lamentation, as I have described, the cry resulted in sociability. But there is no evidence that the cry was stylized, even if it was induced by neocortical reasoning in the young chimp, or that the mother reasoned about it as such. The mother appears to have treated the cry as a simple indexical signal.

Yet the seeds are here for the manipulation of crying into a stylized form. Suppose, for example, that the mother observed the absence of a situation of danger when the child emitted the distress signal. This is, after all, what Goodall herself observed. And chimps seem capable of strategic reasoning about the signals they themselves produce, even if not about the signals they receive. Instead of treating the signal as a simple index, the mother might treat it—if she, in fact, appreciates that no danger is present—as a kind of metasignal. She would recognize the iconic similarity between the vocalization and the danger cry, but she would also recognize its difference, namely, the absence of the external danger. In the situation described by Goodall, Fifi provided no evidence that she made the distinction between the signal and the metasignal. She treated the two in apparently the same way. But the differentiation would seem to be something not far beyond her grasp or that of any intelligent chimp.

Were the differentiation made, the mother would find herself in a position to try out different responses to the son in an effort to solve the problem that the danger-like signal poses for her, namely, the problem of how to get the son to stop emitting the cry, which might potentially disrupt her ongoing activities as well as those of other individuals around her. If she did reason from the relationship between metasignal and signal, she might respond with actions on her part that would resemble her response to the signal but simultaneously be different from them.

Were such the case, for example, in the anecdote above, Fifi might not have taken Frodo up and carried him off but, rather, have interacted with him in some other way that was satisfying to him. A vocal response to the cry on her part, for example, a muffled or partially inhibited cry, might suffice. Frodo, for his part, might also observe that the response was formally distinct from what he would have expected from Fifi through an emission of the danger signal. This could lead him to explore possible modulations of the danger cry that might induce the satisfying vocal response, rather than the "taking up" response associated with the danger cry.

Mutual treatment of crying as metasignal rather than signal seems a giant step forward in the evolution toward language, one that has to be better understood, but it seems also a step that could have taken place long before

language as we know it, or even a minimally contrastive set of syllables, emerged. For once crying operates as metasignal, the larynx and the vocal cords become the targets of neocortical manipulation of an otherwise instinctively controlled vocal apparatus. The transition from instinct to culture, in this area, occurs in the transition from signal to metasignal.

I have attempted to diagram this in Figure 4. There we see (in the upper half) the child's signal—a danger cry in Frodo's case. At the level of a signal, this would be understood (in the absence of any other evidence to the contrary) as a limbically driven, instinctive verbal cry. However, the additional factor observed by Goodall is the absence of a stimulus for the instinctive cry—that is, the absence of a situation of danger. Consequently, Goodall interprets this as a deception on the child's part. The "danger cry," in this situation, is a neocortically fashioned strategic ploy to get the mother's attention. The attention is gotten by making the mother feel that danger might be lurking. If the mother recognized the "danger cry" as a metasignal, however, she might treat it as a problem to be solved, and she might stumble on a communicative solution, namely, producing a signal that sounded like the one her son was producing but was obviously different from it, perhaps a muffled version of the cry, as in Figan's attempt to suppress the food call. Both individuals would be producing

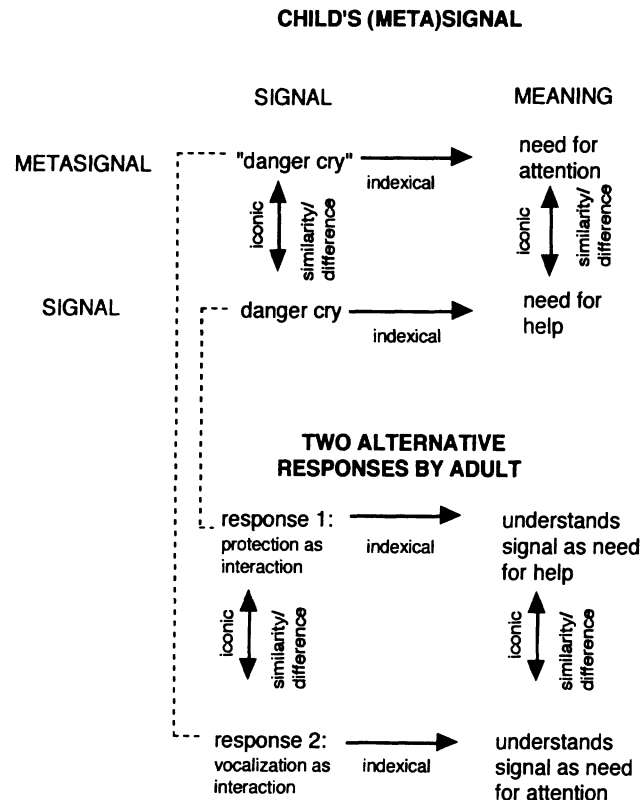


FIGURE 4. Interpretation of a signal as a metasignal versus as a signal. In the two alternative responses to the metasignal, the first response calls attention to the signal, whereas the second calls attention to the metasignal.

something like the same vocal communication. A true cultural signal, neocortically driven, might thus take shape.

The child might have motivation for differentiating the deceptive “danger cry” from a real one. He may want to produce in his mother the verbal response rather than the protective response. The metasignal would become a stylized version of the signal. And it would be neocortically driven (having its origins in strategic manipulation) rather than limbically driven. Hence, it would represent a socially constructed signal that would be, simultaneously, an elementary form of culture.

Note the premium that is placed, in this scenario, on control by the neocortex over the larynx. If there is motivation to stylize cries—whether cries of danger, pain, hunger, or some other sort—then there is motivation to control or shape those cries, and that means control over the larynx. Some control is already present in Frodo’s deception of his mother—if that is, in fact, a correct interpretation of the situation. To deceive, he had to be able to induce the cry. But that induction might be of a primitive sort—a neocortically triggered vocal call but one that was not neocortically nuanced and shaped. Once triggered, it would be indistinguishable from the instinctive call. Neocortical shaping would require attention to more fine-grained modulation of the call, perhaps a partial inhibition of the innate signal. That modulation in turn would come to signal the very desire for sociability itself, a desire to produce the socially appropriate form of the metasignal.

Of course, the scenario I have described ends up looking like stylized or ritualized crying in human populations. In human society, ritual crying does not induce other people to comfort the crier. This is what one would expect if the others were interpreting the cry as a signal, as diagrammed in Figure 4. Instead, ritual crying induces others to ritually cry, as mentioned earlier. People indicate their empathy for the crier by crying themselves.¹¹ And they do so in modulated cries that sound like those of the original lamenter. That stylized lamentation is so widespread in human societies indicates just how successful it is at accomplishing its principal goal—producing sociability.

HOW DO NEOCORTICALLY FASHIONED VOCAL SIGNALS BECOME SOCIALLY STANDARDIZED AND, HENCE, PART OF CULTURE? (THROUGH THE PARENT-CHILD BOND)

Social standardization of the signal within the mother-child relationship is already part of the previous hypothetical account. If the mother responds to the child’s false cry with her own vocalization (modulated cry), the child may learn to modulate its own cry after the image of the mother’s in order to induce the vocal response. To understand more about social standardization, I turn to child language acquisition research to see how it takes place in contemporary humans. Lamentably, the vast bulk of this research in the last few decades focuses on the period of age 18 months and older, that is, the period when the

Saussurean distributional system of language is already apparent. But control over the larynx to produce syllables, in fact, occurs much earlier (MacNeilage and Davis 2000; Vihman 1996; Vihman and DePaolis 2000). Primitive syllables are already taking shape by age four months. At that time, my own daughter was producing a reasonably distinct “gugugu” sound, with the velar fricative [g]-like sound at least somewhat distinguishable from the obscure central vowel. By six months she was making recognizable voiced bilabial [m] sounds. By seven months, she could do bilabial trills, what we commonly call “blowing raspberries.” By eight months, she was making an inventory of sounds her parents distinguished as [p]s, [b]s, [m]s, and [w]s.

What most struck me at the time about her acquisition of linguistic sounds was the way they seemed to take shape progressively out of the crying that characterized her early infancy, as if she were struggling mightily to control that cry, bring it under the sway of the neocortex. At seven months, she produced cry breaks as part of a manipulated stream of sound, where she was learning to take control of the larynx. Her [m]s at seven months often showed tonal modulation and were part of larger cry-like utterances.

In all of this, neocortical control over the vocal cords and larynx was developing in connection with manipulations of the cry. That control emerged out of a great struggle in which she seemed to have been engaged. When she was first born, her wails were spontaneous. By six months she was making, seemingly, great strides toward controlling them and turning the protracted cry into a modulated, syllable-like form. Simultaneously, protoconsonants were taking shape that also involved control over the larynx. What is all of this control about?

My contention is that it is about pre-linguistic but nevertheless vocal communication. In M. A. K. Halliday’s (1975, 1984) study of the infant Nigel, he concludes that by eight months Nigel was communicating both gesturally and vocally. In the vocal realm, Halliday distinguishes two communicative functions Nigel was accomplishing by means of sound, one an interactional function (something equivalent to the command “Be with me”) and the other a kind of world observational function (something like, “You there, see or hear that”). His claim is that Nigel accomplished this communication through the modulation of tone—a low falling tone in the case of “Be with me” and mid-falling tone in the case of “See that.” As vocalizations, these two signals appear to be controlled versions of cries, and interestingly, from an evolutionary perspective, the meaning Halliday infers for the low falling tone, “Be with me,” closely parallels the meaning of ritual lamentation as signal (need for connection), the false cry actually emitted by Frodo, and the postulated evolutionary ancestor of language, a modulated cry signaling a desire for sociability.

Although the two signals Halliday describes for Nigel at eight months are ones that many American parents

readily recognize, we cannot be sure that they are culturally specific signals because they do not sufficiently resemble the linguistic signals produced by adults. Evidence for neocortically controlled, linguistically specific, and socially standardized shapes serving communicative purposes seems to develop between eight months and one year. At 11 months, my daughter was employing a number of recognizable wordlike signals and using them in appropriate contexts, for example, [kikæ] when she noticed a cat going by, the sound shape clearly resembling the parents' regular usage of "kitty cat" in this context, or [dædi] for *daddy*. In Halliday's study, Nigel, when he was one year old, had 28 distinguishable meanings communicated by separate vocalizations. Of these, eight or nine (see Figure 5) appear to be imitations of the adult vocalizations to which Nigel was exposed.¹² Further study by Vihman and DePaolis (2000) confirms the centrality of imitation prior to one year of age.

However, neocortical control over vocalizations at this juncture is not simply a function of imitation of adult forms. Halliday remarks that, during this phase, "there is no obvious source for the great majority of the child's expressions" (1975:22). Yet Nigel was coming up with vocalizations in an attempt to communicate, and his communications were, presumably, in considerable measure neocortically driven. So there are at least two separable components to the achievement of neocortical control over the larynx: one is the strategic deployment of vocalizations in an effort to communicate by differentiating out new vocalizations; the other is the attempt to copy adult forms for communicative purposes. It is that latter copying process that results in standardization.

Social standardization of the signal (and, hence, sharing) depends on the mutual metasignaling discussed in the previous section and summed up in Figure 4. It is not inconceivable that the first reciprocal signal was a muffled or inhibited cry in imitation of the child's vocalization, as I have proposed. In any case, if, by genius or luck, the mother hit on a cry that was obviously iconic with the

child's cry but equally obviously different from it, and if it was a cry that, for whatever reasons, could be readily subjected by the child or others to reasoned interpretation, this might have been the birth of the first socially standardized, neocortically induced vocalization. Hence, it might have been the earliest ancestor of human language.

HOW DO NEOCORTICALLY FASHIONED VOCAL SIGNALS BECOME SOCIALLY STANDARDIZED AND, HENCE, PART OF CULTURE? (THROUGH THE ADULT-ADULT INTERACTIONS)

Social standardization of signal form and meaning within a single mother-child bond would not be much of a precursor to language. Even if it were passed on by the child to its own children and so forth, what evolutionary significance could it have, unless, of course, there were a bottleneck, with only these children surviving? It would seem likely, however, that social standardization of vocalizations would require not just parent-child vocal interactions but also adult-adult vocal interactions. If the signal indicating a desire for social connection were cleverly designed, others might take it up. Although perhaps only Mike profited from his invention of the kerosene can signal, it is possible to imagine an invented vocalization whose ingenious design allowed it to be readily decoded and reproduced by others.

What adaptive significance might it have? If the signal allowed protohumans to solve some of their social problems, without forcing them to disrupt their productive activities, it could be of considerable significance. This is already apparent in the mother-child case. If the mother were not required to stop her productive activity to pick up a child in need of social connection but, rather, could satisfy the child's need vocally while simultaneously continuing to work, some incremental adaptive advantage is gained. Correspondingly, if adults find themselves able to solve some of their interactional problems through such vocalization, without having to halt other activities, they acquire a slight advantage over others.

The development of strategically deployed integrative signals is by no means farfetched. Communicative signals of this sort, albeit not vocal ones, originated in at least one other closely related nonhuman primate group: the bonobos, who deploy sex as a device for easing social tensions within the group (de Waal and Lanting 1997). An invention of the sort I have described, while of immense significance, is not so far removed from inventions that have taken place among our nonhuman primate relatives.

In the case of vocalizations, however, in contrast to the use of sex as a signal, the effect would be to select for neocortical control over the larynx and vocal apparatus. Selective pressure would favor individuals who could not only control their own larynx and vocal apparatus in the service of strategic reasoning but also interpret the vocal signals emitted by others through reasoning about them.

Nigel	Adult	Gloss
[yĩ]	[yes], etc.	"yes I want [object present]"
[ã ă]	[ˈã ă]	"no don't (let's) do that," cf. "uh uh"
[dɛə]	[ðɛə]	"let's look at this together," i.e., "there"
[an:a]	[ana]	"Anna"
[dada]	[dadi], [dædi]	"daddy"
[da]	[dɔg]	"dog"
[ba]	[bas]	"bus"
[pʷi]	[plei]	"let me play with the cat"

FIGURE 5. At 12 to 13.5 months, Nigel had some 28 distinct and communicatively meaningful vocalizations. The eight listed above seem likely imitations of the adult vocalizations. From Halliday 1975:150–151, 1984:25.

The net effect would be individuals primed to invent and interpret new vocalizations.¹³

WHY AND HOW DO SOCIALLY STANDARDIZED VOCAL SIGNALS PROLIFERATE?

Starting with an original ancestral neocortically shaped vocalization signaling the desire for social connection, we are left with the problem of why there should ever have developed more than one signal. Why should new signals arise? The problem is a more general one than imagining specific inventions. We need to understand the processes that spawn new signals.

The processes of creation—really nothing other than I have been describing all along for the original phenomenon of metasignaling—are summed up in Figure 6. To create a new signal, like kerosene can rolling or stylized crying, one differentiates the new signal shape from the old one. If the new signal is to be readily interpreted by others, it must have two important formal properties. First, the new signal must look (or sound, or taste, etc.) sufficiently like the old one that the meaning of the old one can serve as the basis for guessing the meaning of the new one. However, second, the new signal must be obviously and unmistakably distinct from the old one, so that it is recognized as new and, hence, as requiring reasoning to figure out its meaning.

This process is a risky one and highly susceptible to failure, for the signals individuals come up with may be inadequate in one or the other (or even both) ways described above. In Nigel's case, recall that fewer than one-third of his vocalizations resembled adult forms. Although Halliday was able to assign meanings to them, such assignments are by no means straightforward. Children at this age (around one year) are far from perfect communicators. I think we have to imagine that our earliest ancestors—struggling as they must have been to deploy signals that could be interpreted—were probably also far from perfect communicators. We can imagine them trying out new vocalizations, only to find that they were unable to achieve the effect they had contemplated.

Still, the Mikes among them would find success, and gradually the repertoire of shared, socially learned signals would grow. And what a great adaptive advantage this

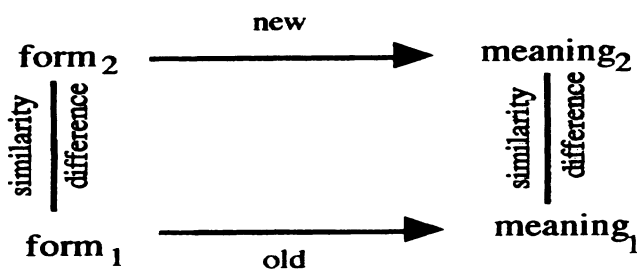


FIGURE 6. Schema for the creation of new signals. The analogical pathway of reasoning described by historical linguists is analogous to the metasignaling pathway proposed here as the earliest precursor of modern linguistic communication.

would be. The ability to neocortically respond to the signals produced by others enables greater flexibility in adaptation. New signals could be created, for example, for warning of new kinds of dangers lurking in novel environments or for the communication of information about edible plants in unfamiliar places, thereby bypassing the operation of biological selection and individual learning. New vocalizations could also be invented to solve novel interpersonal problems arising out of new lifeways, thereby enabling groups of individuals to more effectively coordinate their interactions in the pursuit of collective goals. In thinking about the significance of neocortical reasoning in the production and interpretation of vocalizations, one ends up reciting the litany of benefits associated with culture more generally. Indeed, the solution to the puzzle of neocortically produced and interpreted vocalizations is probably closely tied to the quantitative expansion of the role of culture in the lives of individuals.

So that there is no possible confusion, let me emphasize that the new communicative capability would have been cultural. We are not dealing with one signal being favored over another through biological selection—that is, as the result of genetic mutation producing a different, better adapted limbically triggered signal. Rather, biological selection would favor the capacity to produce and interpret new signals. It would operate to reinforce the connections between neocortex and larynx, rather than between limbic region and larynx. It would produce, in short, a more flexible signaling system, one whose signal forms and meanings could be socially learned and socially transmitted, one in which new signals could be not only produced but also more or less correctly interpreted by others.

We have no fossils of ancestral chimpanzees, presumably because they lived in tropical areas that were not conducive to the fossilization process (Johanson and Edey 1981:363). If the line that led to chimps and the line that led to humans were once indistinguishable, then, or so the reasoning goes, our own earliest ancestors probably lived in that same nonfossilizing area. Therefore, even the very earliest Australopithecine fossils may indicate radiation out of ancestral environments and into new ones.

It is not inconceivable that radiation was made possible by the establishment of socially standardized, neocortically controlled vocal signals.¹⁴ Such signals may have been a leading rather than trailing edge of adaptation. They could thus conceivably have enabled our earliest ancestors to more effectively cope with new and changing environments, beginning a process of radiation that would result in their dispersal around the planet and now, perhaps, beyond.

HOW DO SOCIALLY STANDARDIZED SIGNALS BECOME DETACHED FROM THEIR OBJECTS?

A central mystery of language-origins research is, “How did signs develop that are not directly linked to their objects by contiguity (indexicality) or similarity (iconicity)?”

Studdert-Kennedy describes this detachment as “freedom from control by identifiable external stimuli (displaced reference)” (2000:161). Present-day human languages can be readily deployed to talk about events, objects, people, and places far removed in space and time from the act of speaking, and the signs used to talk about such displaced referents have no detectable physical similarity to the referents themselves. The phrase “the bronzed rocks and the dark jungle” need not be uttered or written near the rocks or jungle it describes, and, in any event, the graphic symbols or spoken words do not physically resemble what they mean. How did such semiotic devices come into existence?

A merit of the metasignaling approach is that it provides an answer to that question. The pathway of reasoning through which the metasignal is produced or interpreted goes through both icons and indexes, as portrayed in Figures 1 and 4. However, the iconic connection is between metasignal and signal, not between signal and object. In other iconic theories of language origins—for example, the onomatopoeic or “bow wow” theory proposed by philologist Max Müller in the 19th century or the gestural origins theories going back to the 18th-century French philosopher Condillac (Hewes 1973)—the iconic connection is between audible sound or visible gesture and whatever it is the sound or gesture represents, for example, the barking of a dog. In the approach proposed here, the iconic connection is between signs and other signs as things in the world. One sign points to another and thereby provides a pathway for reasoning about the relationship between sign vehicles and the world.

A trace of this conception can be found in Hockett’s (1960; Hockett and Ascher 1964) investigations of the design features of language. Hockett focuses on the combinatorial property of language as the one most in need of explaining—the one whose explanation would be key to all the rest. Combination is what makes possible the production of such complex signs as “the bronzed rocks and the dark jungle.” Hockett proposes that the truly major step in language evolution was the production of primitive signal combinations. In his example, which he notes “has never been observed” in modern-day gibbons, the “gibbon finds himself in a situation characterized by both the presence of food and the imminence of danger. The factors are closely balanced. Instead of emitting either the clear food call or the unmistakable danger call, he utters a cry that has some of the characteristics of each” (Hockett and Ascher 1964:142).

In the metasignaling approach, it is not signal combination that is the key development but, rather, signal differentiation. The metasignal is not the result of combining two distinct signals; it is the result of manipulating or modifying an old one. Instead of two limbic stimuli equally balanced, the neocortex seizes just enough control over a limbic signal to mark that signal as distinct. This makes the metasignal salient with respect to the limbic signal but not so salient as to obscure the original signal altogether. As a result, a new pathway of reasoning is

opened up. It runs between signals as well as between the individual sign vehicles and their meanings. Combination would only gradually emerge when the differentiating features became distinguishable as independent signs in their own right.

If this view is correct, then detachment or displacement never really did occur, at least not in a primeval event. Rather, what happened was the gradual insinuation of sign-to-sign pathways of reasoning into the processes of construing the relationship between sign vehicles and their objects. Sign-to-sign relationships produced a new type of meaning, which Deacon (1997), following semiotician Charles Sanders Peirce, refers to as “symbolic,” in opposition (or, perhaps, addition) to indexical and iconic meanings.

Were this neocortical reasoning about signs through metasignals to be of adaptive advantage, pathways of neocortical control over the larynx would be selected for, resulting in the development of the missing element of neuroanatomy reported by Jürgens (2001). The brain would evolve under the influence of culture—that is, social learning—rather than undergoing an independent evolution that made language possible. This echoes the position articulated by Clifford Geertz several decades ago: “Our central nervous system—and most particularly its crowning curse and glory, the neocortex—grew up in great part in interaction with culture” (1973:49).

WHY DO SOCIALLY STANDARDIZED VOCAL SIGNALS ORGANIZE THEMSELVES INTO SYSTEMATIC RELATIONS BASED ON SIMILARITY AND DIFFERENCE AND, HENCE, COME TO RESEMBLE LANGUAGE?

The earliest neocortically fashioned, socially learned vocalizations are a far cry, so to speak, from human languages, with their elaborate lexicons and intricately woven grammars. Similarly, Nigel’s and Jessica’s vocalizations were at eight months of age far from anything resembling the language they would later come to possess. Does this early evolutionary scenario shed any light on the end product?

The kind of semiotic reasoning that I have proposed for our human ancestors can be extended to something resembling grammar. In fact, those familiar with historical structural linguistics will recognize these processes—involving reasoning from the relationship between signal forms to the relationship between their meanings, and vice versa—as the core analogical principle of language more generally. In the present evolutionary scenario, it is the repeated application of this neocortically guided reasoning process that gives rise to new signals and, ultimately, to language, including the morphological complexity of words and the syntactic relations among them.

I do not propose to treat this matter in detail here, for it is so well studied already. Still, it may be helpful to look at one example in order to illustrate the continuing operation of those processes that may have given birth to the earliest metasignals. Take the meaning of the word *alcoholic*—“one who drinks alcohol to excess.” Native speakers

of American English appear to have reasoned about this meaning as if it were meaning₂ in Figure 6 with "alcohol" as meaning₁. They treat the word *alcoholic* as form₂ with respect to *alcohol* as form₁. Because the two meanings are related to each other as "one who drinks alcohol to excess" is to "alcohol," they construe the word form *alcoholic* as containing a formal mark for "one who does _____ to excess."

The question is, What is that formal mark? Were the two signs (*alcohol* and *alcoholic*) the only two in play, the question would be moot. However, the question becomes relevant when the signs form part of a larger set in which one can reason from similarities and differences in physical forms to similarities and differences in meanings. One needs to know what the formal mark is only if one wants to use that formal mark elsewhere to communicate a similar but different meaning. For example, if one thinks about the physical shapes of the words *alcohol* and *alcoholic*, and one understands their meaningful relationship to be that of a liquid called "alcohol" to "a person who drinks alcohol to excess," then what physical shape would one invent to pair with *work* in order to communicate the meaning "one who works to excess"?

In asking that question, one moves from simple meta-signal-signal relationships to complex analogical relationships. It is as if one has two versions of Figure 6, one with the signal *alcohol* and the other with the signal *work*, and one treats the second figure as a metasignal with respect to the first. I have attempted to diagram this in Figure 7. The new signal *X* in Figure 7, in order to convey the new meaning "one who works to excess," must be similar to but different from *work*, and it must also be similar to but different from *alcoholic*. The [ic] portion of *alcoholic* would do, producing the form *workic*. However, because [-ic] is a common word ending in English with its own meaning ("of or related to, associated with, or characterized by"), the similarity to *alcoholic* gets lost. *Workic* would not remind the reader of *alcoholic* in particular. It would rather link *work* to the whole class of which *alcoholic* is one part. For this reason, speakers use the form *-aholic* or *-oholic* together with the form *work*. *Workaholic* is obviously similar to but different from *work*, and, moreover, it is also obviously similar to but different from *alcoholic*. The connection to each can be readily discerned by a native English speaker.

Once the form *-aholic* comes to communicate the meaning "one who does something to excess" in one instance, it can be employed to create an indefinite number of new words in the same way: *chocaholic*, *saltaholic*, *sleepaholic*, *sexaholic*, *TVaholic*, *bookaholic*, *cokeaholic*, and even *junkfoodaholic*.¹⁵ Application of analogy to the basic metasignal-signal principle results in new morphemes and, for this reason, in the phenomenon of sign combination so central to the origins of language in Hockett's construal.

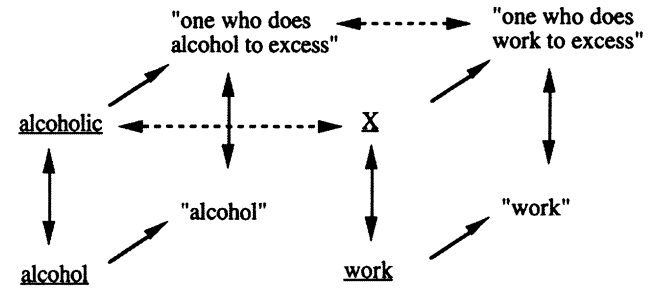


FIGURE 7. The work of analogy in the creation of new signal forms. The question here is: What should *X* be? As a physical form, it should resemble *work*, the way *alcoholic* resembles *alcohol*. In addition, however, it should resemble *alcoholic*. Use of the simple *-ic* form might do if it were not for the fact that it is a widely used ending and hence would not capture the similarity to *alcoholic* in particular. Hence, we have the choice of *-oholic* or *-aholic* as the physical mark carrying the meaning "one who does something to excess."

Analogy has its roots, or so I am proposing, in the most elementary forms of reasoning about the relationship between neocortically induced and innately triggered vocalizations. Once the neocortex is brought into the production and interpretation of vocalizations, proportional reasoning takes hold to generate and interpret new signals. This process leads, inescapably, to the language-like structures of morphology and syntax, with difference operating most prominently in the case of lexical relations and similarity, in the case of grammar.

All of this, in some sense, is contained in the initial semiotic recognitions, however dim, associated with the meta-signal-signal relationship, whereby the metasignal carries over something from the signal through iconicity but asserts its distinctiveness through difference. Distinctiveness indicates the presence of new meaning, but similarity allows a guess as to where that new meaning might be found. The metasignal-signal relationship depends on the coordination of two types of relationship: one between like terms (form to form or meaning to meaning) and the other between unlike terms (form to meaning or meaning to form). And it is this ancestral reasoning about signals, as part of the lifeblood of individuals and communities struggling to adapt to new environments and changing circumstances, that leads ineluctably to the fossilization process to which Emerson alluded. Yes, language is fossil poetry. It is the deposition in stony outline of potent communicative forms invented and deployed by creatures struggling to achieve neocortical mastery over a mercurial world. Should we prove lucky, we may one day, through the mutual reinforcement of distinct lines of research—such as discourse-centered ethnography and comparative neurobiology—catch a glimpse of that pulsating if transient life behind these curious stones.

NOTES

Acknowledgments. A hearty thank you goes to participants in the Penn Humanities Forum Colloquium (October 3, 2000) and the Penn Anthropology Department Colloquium Series on Language Origins (October 5, 2000), where preliminary versions of this article were read and vigorously discussed. I am indebted also to Steve Feld, Marc Meyer, Brandy O'Neil, and Tom Schoenemann for their comments and to Uwe Jürgens for supplying additional information. Last, I consider myself fortunate to be able to count on two able research assistants, Corinne Beiersdorfer and Sara Cott, who helped with manuscript preparation. I thank you all.

Audio examples of ritual lamentation, along with samples of other recorded materials analyzed for this article, can be found at the website <http://www.sas.upenn.edu/~gurban/metasignal/metasignal.htm>.

1. The last two decades have seen a substantial accumulation of research on stylized lamentation. I did a comparative study of central Brazilian lamentational forms, which are sometimes grouped under the heading "welcome of tears" in recognition of their occurrence in connection with the arrival of guests (Graham 1984, 1986; Urban 1988, 1991). Charles Briggs (1992, 1993) has produced some fine-grained studies of lamentation among the Warao of the Orinoco Delta in Venezuela. Farther afield, Steve Feld (1990) published what has since become a classic study of ritualized lament among the Kaluli of New Guinea, for which a recording is also available. Joel Kuipers (1990, 1998) describes laments among the Weyewa of Indonesia. And James Wilce (1998, 1999) looks at the role and form of lament and ritualized complaint in rural Bangladesh.

2. In his late work on the expression of emotions, Darwin (1897) discusses the interplay in humans among the will, habit, and instinct. His concern was to trace the will and habit back to instinct and selection. My concern here is, in some ways, the complementary one: to trace instinct forward to will and habit, that is, strategic reasoning and culture. My central question is: How does the neocortex, rather than instinct, come to control some significant portion of the signaling behavior in humans?

3. Or they express some related meaning, such as welcoming someone who has been away for a long time. Other secondary functions can be added to this basic function of lamentation, for example, criticism of the behavior of others (Briggs 1992).

4. Ritual lament is by no means the only stylized form that communicates by means of metasignal-signal relationships. Formalized dialogues (Urban 1991:123-147) operate in just this way, as probably do the aggressive greetings found in many South American Indian societies. In the contemporary United States, stylized laughter and faked orgasms along with other imitated sexual vocalizations, employed, for example, in the phone sex industry, provide additional examples.

5. By the 1960s, researchers had given up on trying to get chimpanzees to speak English (Hayes 1951; Hayes and Hayes 1952), concluding that the chimps could not gain the necessary control over their vocal apparatus. Building on the obvious manual dexterity of chimps, Allen and Beatrice Gardner (1978) initiated a new line of investigation, using sign language instead of spoken English. Washoe, the Gardner's subject chimp, was able to learn more than 150 discrete signs, and a gorilla, taught by Francine Patterson (1978), learned many more, although the actual numbers in both cases have been disputed (Petitto and Seidenberg 1979; Pinker 1994). Using a computer keyboard, instead of sign language, Sue Savage-Rumbaugh (1994) was able to achieve still more spectacular results with the bonobo Kanzi. It now appears plausible that non-human primates are able to acquire, under human tutelage, at least 400-500 discrete signs. In what measure this results in true language remains in doubt (see Hill 1978; Wallman 1992), but there can be little question that these animals are capable of acquiring (via social learning) and using many more discrete signs than are deployed by nonhuman primates in the wild.

6. Philip Lieberman, in his most recent book (2000), summarizes a vast array of research concerning the neurological bases of language. His concern is to show that language, as a functional neural system, involves "many neuroanatomical structures, many of which also play a part in regulating other aspects of behavior"

(2000:121). The discussions that are most pertinent to the current argument concern voice onset time in syllables. My contention is that the neocortex is central to the ability of humans to control the larynx and, hence, to speak. While not contradicting this claim, Lieberman points to the involvement as well of subcortical structures (basal ganglia) in regulating voice onset time. The key point for our present purposes is that the neocortex does not act alone in turning the larynx on and off but, rather, in concert with other parts of the central nervous system.

7. Precisely how the neocortical signal in the squirrel monkey is relayed to the larynx remains a mystery. Uwe Jürgens, in a personal communication on September 20, 2001, suggested three possible relays: "(1) the small-celled part of the reticular formation of the medulla, (2) the solitary tract nucleus, and (3) the lateral parabrachial region in the dorsal pons."

8. For an enlightening characterization of the relationship between human language and other animal and human forms of semiosis, see Agha 1997.

9. As Terrence Deacon observes, "Despite the fact that cortical motor damage does not disrupt call production in the monkeys that have been studied, motor cortical areas may nevertheless play an indirect role. Projections to oral and vocal motor nuclei in the brain stem may offer a route for direct intentional inhibition of calls" (1997:245).

10. Musicologists use the term *modulate* in the sense of changing keys through regular chord progressions. I use the term here rather in the sense of changing or varying the pitch, amplitude, and voice qualities (for example, creaky versus noncreaky voice).

11. Although a diversion from the present argument, it is nevertheless interesting to consider, in this context, the role of tears in crying. Darwin observed the absence of tears in the crying of very young infants:

Infants whilst young do not shed tears or weep, as is well known to nurses and medical men. This circumstance is not exclusively due to the lachrymal glands being as yet incapable of secreting tears. I first noticed this fact from having accidentally brushed with the cuff of my coat the open eye of one of my infants, when seventy-seven days old, causing this eye to water freely; and though the child screamed violently, the other eye remained dry, or was only slightly suffused with tears. A similar slight effusion occurred ten days previously in both eyes during a screaming-fit. The tears did not run over the eyelids and roll down the cheeks of this child, whilst screaming badly, when 122 days old. This first happened 17 days later, at the age of 139 days. [1897:152]

More recent research (Frey and Langseth 1985; van Haeringen 1981; cf. Lutz 1999), while still inconclusive, has suggested a possible difference in the chemical composition of reflex tears (produced by an irritant such as onion vapors in the eye) and those produced through some kind of neocortical involvement (for example, in response to watching a tearjerker movie). Can tears associated with weeping be an effect of neocortical involvement? Are tears the products, in infancy, of progressive neocortical control over cries? If that control is at first blunt, rather than fine grained, then perhaps a range of systems, including those that stimulate the lachrymal glands, is being affected. As more fine-grained control is achieved, once the "lament" as metasignal has differentiated into more specific communicative signals, the original lament metasignal can still trigger lachrymal involvement.

12. Halliday's approach to language is functional. People, including children, accomplish tasks by means of language deployment, so that the doing (or what I called earlier "lifeblood") of language is emphasized over the abstract form (or fossil, in Emerson's sense). This view is akin to that espoused by P. Thomas Schoenemann (1999), who regards the development of language as a response to the need to communicate. My own concern here, while related to these, is, rather, with how neocortical control over vocalization might have developed and how that control would go along with the production of truly cultural signals and meanings.

13. Darwin (1871) and others (Deacon 1997; Lieberman 2000) have hypothesized the origins of language in the musical interactions between males and females. Courtship songs, to be sure, have

developed in other species—but as part of innately triggered vocalizations. Although courtship songs may have arisen in the early evolution of the human line, if they were spawned by the kinds of processes hypothesized here for mother–child interactions, they too would have to be strategically reasoned inventions, like Mike’s kerosene cans, built on the model of already existing signals but distinguished from them in obvious and important ways. As neocortically controlled signals, they could then be socially transmitted from individual to individual as part of culture. The advantage here would be the higher probability of sexual selection, with cleverly constructed vocal signals resulting in increased likelihood of mating. Music as the precursor to language—a notion not so far removed from that of modulated cries, such as I have proposed—is an old idea in the philosophical literature, being especially prominent in the 18th century in the writings of Rousseau and Herder, among others (Tomlinson 2000).

14. One should not conclude from this that language as we know it today must have emerged at some remote period. The elementary communicative forms I am suggesting are very far, in evolutionary terms, from modern language. It is not inconceivable, indeed, that the full-blown form of language did not emerge until the dispersion of anatomically modern humans in the past 200,000 years or even, for that matter, the 30–40,000-year time frame proposed by Chase and Dibble (1987).

15. I have attested all of these forms in the speech of university students. Some have not and may never achieve wide circulation, but the possibility is there through the work of analogy.

REFERENCES CITED

- Agha, Asif
1997 “Concept” and “Communication” in Evolutionary Terms. *Semiotica* 116(2–4):189–215.
- Briggs, Charles
1992 “Since I Am a Woman, I Will Chastise My Relatives”: Gender, Reported Speech, and the (Re)Production of Social Relations in Warao Ritual Wailing. *American Ethnologist* 19:337–361.
1993 Personal Sentiments and Polyphonic Voices in Warao Women’s Ritual Wailing: Music and Poetics in a Critical and Collective Discourse. *American Anthropologist* 95:929–957.
- Chase, Philip, and Harold Dibble
1987 Symbolic Behavior in the Middle Paleolithic: A Review of Current Evidence and Interpretations. *Journal of Anthropological Archaeology* 6:263–296.
- Darwin, Charles
1871 *The Descent of Man*. London: John Murray.
1897 *The Expression of the Emotions in Man and Animals*. New York: D. Appleton and Co.
- Deacon, Terrence W.
1997 *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: W. W. Norton and Co.
- de Waal, Frans, and Frans Lanting
1997 *Bonobo: The Forgotten Ape*. Berkeley: University of California Press.
- Feld, Steven
1990 *Sound and Sentiment: Birds, Weeping, Poetics, and Song in Kaluli Expression*. Philadelphia: University of Pennsylvania Press.
- Frey, William H., and Muriel Langseth
1985 *Crying: The Mystery of Tears*. New York: Harper and Row.
- Gardner, Allen, and Beatrice T. Gardner
1978 Comparative Psychology and Language Acquisition. *Annals of the New York Academy of Sciences* 309:37–76.
- Geertz, Clifford
1973 *The Interpretation of Cultures*. New York: Basic Books.
- Goodall, Jane
1968 Behaviour of Free-Living Chimpanzees of the Gombe Stream Area. *Animal Behaviour Monographs* 1:163–311.
1986 *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, MA: Harvard University Press.
- Graham, Laura
1984 Semanticity and Melody: Parameters of Contrast in Shavante Vocal Expression. *The Latin American Music Review* 5:161–185.
1986 Three Modes of Shavante Vocal Expression: Wailing, Collective Singing, and Political Oratory. *In Native South American Dis-*
- course. Joel Sherzer and Greg Urban, eds. Pp. 83–118. Berlin: Mouton de Gruyter.
- Halliday, M. A. K.
1975 *Learning How to Mean: Explorations in the Development of Language*. Amsterdam: Elsevier.
1984 Language as Code and Language as Behaviour: A Systemic-Functional Interpretation of the Nature and Ontogenesis of Dialogue. *In The Semiotics of Culture and Language*, vol. 1: Language as Social Semiotic. Robin P. Fawcett, Michael A. K. Halliday, Sydney M. Lamb, and Adam Makkai, eds. Pp. 3–35. London: Frances Pinter Publishers.
- Hayes, Catherine
1951 *The Ape in Our House*. New York: Harper and Brothers.
- Hayes, Keith, and Catherine Hayes
1952 Imitation in a Home-Raised Chimpanzee. *Journal of Comparative Physiological Psychology* 45:105–109.
- Hewes, Gordon
1973 Primate Communication and the Gestural Origin of Language. *Current Anthropology* 14(1–2):5–24.
- Hill, Jane H.
1978 Apes and Language. *Annual Review of Anthropology* 7:89–112.
- Hockett, Charles F.
1960 The Origin of Language. *Scientific American* 203:88–96.
- Hockett, Charles F., and Robert Ascher
1964 The Human Revolution. *Current Anthropology* 5(3):135–168.
- Jakobson, Roman
1971 Why “Mama” and “Papa”? *In Selected Writings*, vol. 1: Phonological Studies. Pp. 538–545. The Hague: Mouton.
- Jakobson, Roman, and Linda R. Waugh
1987 *The Sound Shape of Language*. Berlin: Mouton de Gruyter.
- Johanson, Donald C., and Maitland A. Edey
1981 *Lucy: The Beginnings of Humankind*. New York: Simon and Schuster.
- Jürgens, Uwe
2001 A Comparison of the Neural Systems Underlying Speech and Nonspeech Vocal Utterances. Electronic document, <http://baserv.uci.kun.nl/los/Articles/jrgens.html>, accessed September 5.
- Jürgens, Uwe, and P. Zwirner
1996 The Role of the Periaqueductal Grey in Limbic and Neocortical Vocal Fold Control. *Neuroreport* 7(18):2921–2923.
- Kuipers, Joel C.
1990 Power in Performance: The Creation of Textual Authority in Weyewa Ritual Speech. Philadelphia: University of Pennsylvania Press.
1998 Language, Marginality and Identity in Indonesia: The Shifting Place of Ritual Speech on Sumba. Cambridge: Cambridge University Press.
- Liberman, Mark
2000 Linguistic Properties of Human Bodies, Brains, Minds and Cultures. Paper presented at “The Origins of Language,” Penn Anthropology Department Colloquium Series 2000–01, Philadelphia, October 19.
- Lieberman, Philip
2000 Human Language and Our Reptilian Brain: The Subcortical Bases of Speech, Syntax, and Thought. Cambridge, MA: Harvard University Press.
- Lutz, Tom
1999 *Crying: A Natural and Cultural History of Tears*. New York: W. W. Norton.
- MacNeilage, Peter F., and Barbara L. Davis
2000 Evolution of Speech: The Relation between Ontogeny and Phylogeny. *In The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Chris Knight, Michael Studdert-Kennedy, and James R. Hurford, eds. Pp. 146–160. Cambridge: Cambridge University Press.
- Patterson, Francine G.
1978 The Gestures of a Gorilla: Language Acquisition in Another Pongid. *Brain and Language* 5:56–71.
- Petitto, L. A., and M. S. Seidenberg
1979 On the Evidence for Linguistic Abilities in Signing Apes. *Brain and Language* 8:162–183.

Pinker, Steven

1994 *The Language Instinct: How the Mind Creates Language*. New York: HarperCollins Publishers.

Savage-Rumbaugh, Sue

1994 *Kanzi: The Ape at the Brink of the Human Mind*. New York: John Wiley and Sons.

Schoenemann, P. Thomas

1999 Syntax as an Emergent Characteristic of the Evolution of Semantic Complexity. *Minds and Machines* 9:309–346.

Studdert-Kennedy, Michael

2000 Evolutionary Implications of the Particulate Principle: Imitation and the Dissociation of Phonetic Form from Semantic Function. *In The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Chris Knight, Michael Studdert-Kennedy, and James R. Hurford, eds. Pp. 161–176. Cambridge: Cambridge University Press.

Tomlinson, Gary

2000 Singing the Same Old Song? View on the Origin of Language, 1600–2000. Paper presented at “The Origins of Language,” Penn Anthropology Department Colloquium Series 2000–01, Philadelphia, September 20.

Urban, Greg

1988 Ritual Wailing in Amerindian Brazil. *American Anthropologist* 90(2):385–400.

1991 *A Discourse-Centered Approach to Culture: Native South American Myths and Rituals*. Austin: University of Texas Press.

van Haeringen, Nicholas J.

1981 Clinical Biochemistry of Tears. *Survey of Ophthalmology* 26(2):84–96.

Vihman, Marilyn M.

1996 *Phonological Development: The Origins of Language in the Child*. Oxford: Blackwell.

Vihman, Marilyn M., and Rory A. DePaolis

2000 The Role of Mimesis in Infant Language Development: Evidence for Phylogeny? *In The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Chris Knight, Michael Studdert-Kennedy, and James R. Hurford, eds. Pp. 130–145. Cambridge: Cambridge University Press.

Wallman, Joel

1992 *Aping Language*. Cambridge: Cambridge University Press.

Wilce, James

1998 *Eloquence in Trouble: The Poetics and Politics of Complaint in Rural Bangladesh*. New York: Oxford University Press.

1999 Transforming Laments: Performativity and Rationalization as Linguistic Ideologies. *In Languages of Sentiment*. Gary B. Palmer and Debra J. Occhi, eds. Pp. 39–64. Amsterdam: John Benjamins.