Knight, C. 2008. Comment on Arbib, M., K. Liebal and S. Pika, Primate Vocalization, gesture and the evolution of human language. *Current Anthropology*, **49** (6): 1064-1065.

Arbib, Liebal, and Pika provide an excellent—and long overdue—comparative survey of the incidence of gestural versus vocal communication in nonhuman primates. I like their proposal that the primate mirror neuron system underpinning gestural imitation played a key role in enabling language parity. I am also persuaded by their more general argument that the emergence of vocal speech in our ancestors in some way presupposed the scaffolding provided by gesture and then pantomime.

Unfortunately, the article ends rather abruptly, having offered little that merits description as an actual theory. The authors address a range of "what," "when," and "how" questions yet never ask the crucial question "Why?" Yes, apes in general do lack volitional control over their vocal signals. Yes, they do seem to have much greater control over their manual gestures. And, yes, manual gestures in humans "can have an effect on vocalizations made at the same time, thus creating certain natural vocal concomitants of manual efforts." It would therefore follow logically that one way an intelligent primate might enhance cognitive control over its vocal signaling would be by intentionally jumping around or otherwise manipulating its body so as to influence any sounds being emitted at the same time.

But all this strikes me as a strangely mechanistic approach to the theoretical difficulties—as if no ape or monkey ever thought to modulate its vocal signals by deploying the equipment it already has. There must surely be some more plausible reason why these animals in fact do not play around creatively or imaginatively with vocal communication. After all, young primates can be strikingly creative and imaginative in their playful antics. In the interests of masticatory efficiency, moreover, they possess jaws, lips, and tongues that are subject to fine motor control. Little effort is needed to activate the relevant mouth muscles. If greater signal flexibility would be adaptive, why not use such ready-made, highly efficient equipment to modulate sounds in the way humans do?

Instead of restricting ourselves to yet another description and

classification of signaling modalities and corresponding mechanisms, we surely need some Darwinian thinking here. Among nonhuman primates, what selection pressures might have rendered it *adaptive* for vocal communication to be so strikingly insulated from cognitive control? What fitness advantages might accrue to an intelligent ape from its *inability* to play around with its vocal signals? Such questions cry out for an answer. If we do not even address them, we are unlikely to get far in elucidating the evolutionary relationships among primate vocalization, primate gesture, and speech-based human language.

The ability to engage in pantomime is, by definition, an ability to fake one's bodily signals and displays. For patent fakes to be accepted as valid currency for purposes of communication, unusually high levels of social cooperation and corresponding trust must be assumed. This presents a theoretical conundrum because those primates intelligent enough to deploy such potentially deceptive strategies will also be clever enough to competitively exploit the trust presupposed by their habitual use (Knight 1998). This could explain why, despite their quite developed capacities for deploying and comprehending symbolic conventions when in captivity, nonhuman primates apparently find so little use for symbolic communication in the wild (Ulbaek 1998).

What would happen if a Machiavellian mutant monkey *did* discover that it could freely substitute one predator alarm call for another, regardless of the presence of any actual threat? Insofar as the fakes were exploited for purposes of tactical deception, they would lose their former status as reliable— hence meaningful—signals. To the extent that salient aspects of any signal can be intentionally faked, conspecifics will simply ignore those variable aspects in favor of any hard-to-fake acoustic features that might prove unintentionally significant. In a Darwinian social world, selection pressures will in this way drive signalers to persuade receivers of the reliability of their signals by demonstrating precisely that they are *not* subject to cognitive control.

This will apply in particular to vocal signaling, which works at a distance, often in contexts that do not allow opportunities for immediate verification. Sound signals go around corners, work in the dark, operate over distances, and leave signalers free to continue with noncommunicative manual tasks. Such advantages make it especially important to protect the vocalauditory modality from deceptive abuse. Lack of volitional control acts like the watermark on a banknote—it

proves that the owner was not the printer. The need to guarantee reliability applies less to visual signals used in face-to-face interactions because such contexts generally offer little scope for abuse.

Facial and manual gesture work best at close quarters, in intimate contexts where immediate verification should be relatively easy. Opportunities for deception are correspondingly few. For example, when one chimpanzee informs a grooming partner at which point on its body it needs to be scratched (Pika and Mitani 2006), what could it possibly gain from a deceptive signal? It is surely no coincidence that nonhuman primates get closest to volitional referential signaling in those restricted social contexts that offer the least scope for deceptive abuse.

But this is precisely the theoretical problem: human language is not used primarily as an aid to ongoing physical activities such as grooming. Its distinctive function is "displaced reference"—communication about things not currently within sensory range. No mechanistic approach of the kind exemplified by Arbib and his colleagues can measure up to the challenge of explaining how this kind of language could possibly have evolved.

References Cited

Knight, C. 1998. Ritual/speech coevolution: A solution to the problem of deception. In *Approaches to the evolution of language: Social and cognitive bases,* ed. J. R. Hurford, M. Studdert-Kennedy, and C. Knight, 68–91. Cambridge: Cambridge University Press.

Pika, S., and J. C. Mitani. 2006. Referential gesturing in wild chimpanzees (Pan troglodytes). *Current Biology* **16**:191–92.

Ulbaek, I. 1998. The origin of language and cognition. In *Approaches to the evolution of language: Social and cognitive bases,* ed. J. R. Hurford, M. Studdert-Kennedy, and C. Knight, 30–43. Cambridge: Cambridge University Press.