

For example, Morford and Goldin-Meadow (1992) discovered that spontaneous hand gestures produced by an adult facilitated one- and two-year-olds' understanding of the adult's accompanying speech. Moreover, Kelly (2001) argued that hand gestures may interact synergistically with speech to help toddlers "break into" an understanding of complex pragmatic communication (e.g., saying "It's almost time for dinner" while pointing to a mess in front of a child). Finally, Kelly et al. (1999) demonstrated that speech and gesture mutually disambiguated one another during adult language comprehension – that is, gesture not only disambiguated the meaning of speech, but *speech itself* disambiguated the meaning of gesture.

Hence, there are solid behavioral data that suggest that speech and gesture are tightly integrated in present-day communication. But what is going under the surface of this behavior? The strongest evidence that speech and gestures are linked during communication comes from recent neuroscience studies investigating how the brain processes language. For example, Rizzolatti and Arbib (1998) theorized that traditional language areas in the human brain (e.g., Broca's area) may be involved in both the processing of language *and* the processing of hand motions. Further support that language and gesture may be linked in the brain comes from Pulvermüller and colleagues (Pulvermüller et al. 2001). They used a high-resolution EEG technique during a verb comprehension task and discovered that comprehension of action verbs activated parts of the primary motor cortex that were physically associated with those verbs (e.g., the verb "catch" activated arm regions, and the verb "kick" activated leg regions). These studies suggest that language and action regions in the brain have a close relationship during production and comprehension. However, no study to date has directly investigated how speech and gestural actions are integrated in the brain during real-time language processing.

Currently, this commentator is using a high-density event-related potential (ERP) technique to investigate this issue (Kelly 2003). This study measured ERPs to speech while adults viewed video segments of people speaking and gesturing about various objects. Preliminary analyses suggest that bilateral frontal sites differentiated speech that was not accompanied by gesture (e.g., saying the word "tall" without gesturing), from speech that was accompanied by matching gesture (e.g., saying the word "tall" while gesturing to a tall, thin object) and mismatching gesture (e.g., saying the word "tall" while gesturing to a short, wide object). Specifically, there was a greater negativity from 320 msec to 600 msec for the no-gesture stimuli compared to the matching and mismatching stimuli. This suggests that the brain processes speech that is accompanied by gesture differently from speech that is not.

The most interesting finding was that ERPs to the speech were different even within the different gesture conditions. Specifically, there was a classic N400 effect in the bilateral temporal regions for the mismatching but not matching stimuli.¹ It is important to note that the speech tokens in both conditions were *identical*, with the only difference between conditions being the different accompanying gestures. These results suggest that the brain integrates gestural information into its processing of speech fairly early in the comprehension process, and provides evidence that gesture and speech are tightly integrated in language processing. This type of research provides vestigial support for Corballis's general argument that speech and gesture were linked in our evolutionary past.

With specific regard to Corballis's lateralization argument, an interesting follow-up to the above study would be to investigate the influence that handedness plays in the brain's *production* of speech and gesture. For example, does the brain process right-handed gestures differently from left-handed gestures in language production? Perhaps by using neuroscience techniques that are relatively resistant to motion artifacts, one could investigate whether right-handed individuals demonstrate different neural patterns of language activation when they produce right-handed

versus left-handed gestures along with speech. If confirmed, this would provide further "present-day" support for Corballis's intriguing argument.

NOTE

1. The N400 effect reflects the unconscious neural integration of semantic information during language comprehension (Kutas & Hillyard 1980).

The secret of lateralisation is trust

Chris Knight

Department of Anthropology and Sociology, University of East London, Dagenham, Essex, RM8 2AS, United Kingdom. c.knight@uel.ac.uk

Abstract: Human right-handedness does not originate in vocalisation as such but in selection pressures for structuring complex sequences of digital signals internally, as if in a vacuum. Cautious receivers cannot automatically accept signals in this way. Biological displays are subjected to contextual scrutiny on a signal-by-signal basis – a task requiring coordination of both hemispheres. In order to explain left cerebral dominance in human manual and vocal signalling, we must therefore ask why it became adaptive for receivers to abandon caution, processing zero-cost signals rapidly and on trust.

My difficulty is with the core of Corballis's argument. Why should exempting the hands from their former communicational responsibilities have had the paradoxical effect of extending left-hemispheric control to these now-excluded hands? Primate-style vocalisations are controlled quite differently from modern speech. The "most critical adaptation necessary for the evolution of speech," as Corballis himself explains (sect 2.4), "was the change in brain organization that resulted in the intentional control of vocalization." Right-handedness is said to have emerged through the hands' involvement with vocal speech – but only as and when vocal signals themselves were becoming as easy to manipulate as manual ones, and only at a very late stage, when manual gesture was in fact being phased out. Presumably, then, during this critical period, specialised brain mechanisms for controlling manual, chewing, and other precisely calibrated sequential movements were extending their remit to previously irrepressible vocalisations. Insofar as these manipulative mechanisms imposed hierarchical order on formerly nonsyntactical vocal sequences, we might plausibly infer that they were already left-lateralised. Yet Corballis's explanation for right-handedness is the reverse of this – anciently left-lateralised centres of vocal control are said to have become extended to govern the hands. It may well be that the apparent contradiction can be resolved, but currently the direction of causality in this argument appears to me quite unclear.

A basic constraint in biological signalling is that if you can intentionally manipulate a signal, then you can fake it. Darwinian signal-evolution theory – not drawn upon by Corballis – sets out from the assumption that, without group-level public sanctions, generalised intentional honesty cannot be sustained. Except in the case of *Homo sapiens*, group-level moral codes are impossible – no biological population can afford to sustain the required system of sanctions. Intentional honesty is therefore an unrealistic assumption for receivers to make. This is why, throughout the entire history of life on earth, no biological species prior to *Homo sapiens* even so much as began to communicate on the basis of a conventional code. Conventional signalling is in this sense like "group selection" – theoretically conceivable but in practice of no Darwinian significance (Zahavi 1993). It does not happen because in a competitive world, no one can afford to remain faithful to the extremely costly contractual understandings and commitments which would have to be assumed.

By contrast, the secret of human left hemispheric specialisation – like the secret of language itself – is trust. Brain lateralisation is

driven by selection pressures to sequence, manipulate, and impose hierarchical order on low-cost digital alternations internally as if in a vacuum. But one side of the brain must be anchored in necessity if the other is to experiment with such freedom. One part of the brain must stay alert if the other is to become lost in its own signals. In just the same way, one foot must bear the weight of the dancer's body if the other is to trace fancy patterns in the air, or one hand must grip the slate if the other is to draw marks across its surface. Where the overall context is purely biological, the freely autonomous – normally left-lateralised – activity of imposing structure can certainly still take place. But the resulting movements will not qualify as socially trustworthy signals, being disqualified precisely for appearing so variable and unconstrained.

Even in nature, however, the songs of songbirds and cetaceans show that low-cost autonomous modulations can play a signalling role – on condition that they occur as variables within an otherwise costly, nonarbitrary, and therefore meaningful display. An example will illustrate this point. A weak or frightened animal is likely to be cautious, tentative, and exploratory. It must alternate between action and reaction, coordinating inputs from both hemispheres as it scans the environment for fresh information in advance of each new decision. Normally, for example, it would be risky for a songbird to shut its eyes or block off its ears. Paradoxically, however, for a babbler to “show off” that it can afford to do just that – to sing as if only the song mattered – can be an impressive display of self-confidence. Zahavi and Zahavi (1997) explain this as follows:

Why do babblers use precisely spaced syllables only when they are eager to fight? In order to emit rhythmic, regularly spaced, and clearly defined syllables, one has to concentrate on the act of calling. Any distraction – such as a glance sideways – distorts both the rhythm and the precision of sound; an individual cannot at one and the same time collect information and concentrate on performance. A call composed of precise, rhythmic syllables testifies that the caller is deliberately depriving itself of information, which means either that it is very sure of itself or that it is very motivated to attack, or both. (p. 21)

The Zahavis add that a human being who is in control of a situation likewise tends to issue threats in an ordered, rhythmic sequence, as if celebrating the fact that external reality can be ignored.

To disconnect from reality is to lose touch with the right brain. Less dominant figures cannot afford to do this, which may explain why they tend to rely more heavily on the right hemisphere while speaking (Armstrong & Katz 1983; Ten Houten 1976). Phonological processing is certainly less lateralised in human females than in males (Shaywitz et al. 1995). Lack of dominance makes it vital to stay sensitive to the total environment, drawing on the right hemisphere in order to do so. But autonomous left hemispheric control does not necessarily imply personal dominance. Its fundamental precondition is simply that low-cost signals – whether manual or vocal – need take no account of environmental feedback or resistance. The confident songbird shows off by “deliberately depriving itself of information,” ceding priority to the left hemisphere in the process. When signals need only connect up with one another, free of any requirement to engage with the external environment, it makes sense to encapsulate the computational circuits close together in one cerebral hemisphere while allowing the other to remain in touch with temporarily irrelevant reality.

Following Kobayashi and Kohshima (2001), Corballis notes that humans differ from primates in that human eyes are not inscrutable but enhance cognitive transparency. But this difference is more than an incidental curiosity. Ancestral social networks – even for sexually mature humans – must have been by primate standards anomalously supportive, making it safe to assume that anyone close enough to see the whites of the eyes was likely to be friend, not foe. Direction of gaze is an aspect of ordinary vision. But it may incidentally serve as a signal. A deliberate “wink” can speak volumes at virtually zero cost. Speech may be conceptu-

alised as an extension of the same principle. Where trust is sufficiently high, resistance on the part of listeners disappears, allowing the subtlest of signals to produce effects. Comprehension now involves inserting oneself imaginatively in the signaller's mind (Tomasello 1999). Speech signals do not need to generate their own trust – at the most basic processing level, an assumption of automatic trust is already built in. In fact, on this level it is legitimate to assume a conflict-free – in Chomsky's (1965, p. 3) terms, “completely homogenous” – speech community. So great is the trust, that language works almost as if one component of the brain – or one component of a computing machine – were simply transmitting digital instructions to another (Chomsky 1995; 2002). Quite regardless of whether signs are manual or vocal, it is this bizarre situation which liberates the potential of one hemisphere to arrange complexity independently of the other. We are left with a puzzling intellectual challenge: to elucidate how the necessary levels of trust could ever have been compatible with our selfish genes. Because I believe this to be the key theoretical issue, it will not surprise Corballis that I am critical of his thought-provoking but non-adaptive account, preferring my own more explicitly Darwinian alternative (Knight 1998; 1999; 2000; 2002).

ACKNOWLEDGMENT

Support from the Leverhulme Trust in the associated research is gratefully acknowledged.

Integration of visual and vocal communication: Evidence for Miocene origins

David A. Leavens

Psychology Group, School of Cognitive and Computing Sciences, University of Sussex, Falmer, East Sussex BN1 9QH, United Kingdom.

davidl@central.susx.ac.uk

<http://www.cogs.susx.ac.uk/users/davidl/index.html>

Abstract: Corballis suggests that apes lack voluntary control over their vocal production. However, recent evidence implicates voluntary control of vocalizations in apes, which suggests that intentional control of vocal communication predates the hominid-pongid split. Furthermore, the ease with which apes in captivity manipulate the visual attention of observers implies a common cognitive basis for joint attention in humans and apes.

Corballis suggests that intentionality in communication is exhibited in the visual domain by many primate species (sect. 2.1), but that voluntary control of vocalizations evolved uniquely within our lineage, sometime after the time when gestural language emerged, possibly as late as several hundred thousand years ago. Corballis states that “chimpanzee calls surely have little, if any, of the voluntary control and flexibility of human speech” (sect. 2.1).

Voluntary control over gestural communication by apes is well established (e.g., Leavens 2001; Leavens et al. 1996; Tomasello & Call 1997; Woodruff & Premack 1979), as Corballis notes (sect. 2.1). No researcher can speak to the state of mind of their ape (or human infant) subjects, but operational criteria for intentional communication are relatively standard and uncontroversial in both comparative psychology (e.g., Leavens & Hopkins 1998) and developmental psychology (Bard 1992). Among other criteria, intentional communication requires an audience and is sensitive to changes in the behavioral cues to attention in the audience. With some few exceptions (e.g., Povinelli & Eddy 1996), virtually all experimental and observational studies have confirmed these operational criteria of intentional communication in the gestural production of both free-ranging and captive apes (e.g., Bard 1992; Call & Tomasello 1994; Hostetter et al. 2001; Krause & Fouts 1997; Leavens et al. 1996; Tomasello et al. 1994).

Evidence is growing which is consistent with the interpretation that some voluntary control over vocal production is exhibited by