



## Essay

## The language void: the need for multimodality in primate communication research

Katie E. Slocombe<sup>a,\*</sup>, Bridget M. Waller<sup>b,1</sup>, Katja Liebal<sup>b,c,d,2</sup>

<sup>a</sup> Department of Psychology, University of York

<sup>b</sup> Department of Psychology, University of Portsmouth

<sup>c</sup> Languages of Emotion Cluster, Free University Berlin

<sup>d</sup> Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

### ARTICLE INFO

#### Article history:

Received 6 May 2010

Initial acceptance 6 July 2010

Final acceptance 25 January 2011

Available online 9 March 2011

MS. number: 10-00311R

#### Keywords:

facial expression

gesture

language evolution

multimodal communication

primate communication

vocalization

Theories of language evolution often draw heavily on comparative evidence of the communicative abilities of extant nonhuman primates (primates). Many theories have argued exclusively for a unimodal origin of language, usually gestural or vocal. Theories are often strengthened by research on primates that indicates the absence of certain linguistic precursors in the opposing communicative modality. However, a systematic review of the primate communication literature reveals that vocal, gestural and facial signals have attracted differing theoretical and methodological approaches, rendering cross-modal comparisons problematic. The validity of the theories based on such comparisons can therefore be questioned. We propose that these a priori biases, inherent in unimodal research, highlight the need for integrated multimodal research. By examining communicative signals in concert we can both avoid methodological discontinuities as well as better understand the phylogenetic precursors to human language as part of a multimodal system.

© 2011 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

It is often proposed that language is one of *the* features separating humankind from other animals and so researchers have long tried to establish how, why and in what ways language is different from other animal communication systems (e.g. Hockett 1960). Language seems to result from a complex interaction of cultural and genetic inheritance (Lachlan & Feldman 2003). Regardless of the specific mechanisms involved, it seems likely, however, that language built on cognitive and communicative skills already present in our primate lineage, rather than evolving from scratch in humans (Hauser et al. 2002). Using the assumption that the skills present in related species would also have been present in the common ancestors of those species, the comparative approach attempts to identify which skills could have served as building blocks for language. As language seems to have evolved uniquely in humans, one cannot explicitly identify components of language in other species; nevertheless, the

search for homologous traits in our primate cousins can identify *potential* precursors to human linguistic abilities.

The research concentrating on commonalities in species closely related to humans is complemented by research on more distantly related species. Convergent evolution of similar traits in distantly related species is important for estimating the selection pressures that favoured the evolution of these traits. Although this line of research makes an essential contribution to the language evolution debate, in this essay we focus on the contribution primate research makes in the search for potential precursors to language with only some exploration of examples in other species.

It is possible that language built on abilities co-opted from cognitive domains indirectly related to communication, such as general mechanisms of learning and memory (e.g. Endress et al. 2009). However, since many primate species use gestures, vocalizations and facial expressions to communicate, most comparative work in this field focuses on communicative behaviour and the cognitive skills underlying communication as potential precursors to our own linguistic abilities. Although language is predominantly expressed through speech in modern humans, sign and written languages show that language can also be expressed through visual modalities. Thus, being a cognitive faculty, language is not modality specific and, as such, could have built on abilities and skills present in any communicative modality. In contrast, to understand the evolution of *speech*

\* Correspondence: K. E. Slocombe, Department of Psychology, University of York, York YO10 5DD, U.K.

E-mail address: [katie.slocombe@york.ac.uk](mailto:katie.slocombe@york.ac.uk) (K. E. Slocombe).

<sup>1</sup> B. M. Waller is at the Department of Psychology, University of Portsmouth, King Henry 1st Street, Portsmouth PO4 0LL, U.K.

<sup>2</sup> K. Liebal is at Evolutionary Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, Berlin 14195, Germany.

within a comparative framework, one must look at vocal/facial communication in other animals.

It is not universally accepted that meaningful similarities between animal signals and human language exist, with some advocating that animal signals are qualitatively different to language on many levels. For example, Bickerton (1992) argued that animal communication is generally holistic rather than componential and lacks grammatical items and the infinite productivity seen in language. Bickerton argued that such differences limit the utility of comparative research for understanding language evolution. However, among the researchers who do pursue the identification of commonalities between language and animal signals there is currently fierce debate between proponents of two seemingly mutually exclusive theoretical approaches: some identify primate gestures as the most promising precursors for human language, while others propose vocalizations as the most promising precursor. In contrast, facial expression is rarely proposed as a linguistic precursor.

#### *Gestural Theories of Language Evolution*

Theories proposing a gestural origin of language suggest that spoken language was preceded by a gestural stage using visual, voluntarily controlled signals (Hewes 1973). Such theories emphasize the similarities between primate gestural communication and human language. First, gestural theories draw on the discovery of mirror neurons and the apparent link between manual gestures and homologous areas of language production in the monkey brain (Rizzolatti & Arbib 1998). Second, it is through gestures, not vocalizations, that attempts to teach apes human language systems, such as American Sign Language, have succeeded with individuals such as Washoe and Koko (Hayes 1951 versus Gardner & Gardner 1969; Patterson 1978). Finally, when examining communication between conspecifics, gestural theories highlight the flexible usage of gestures, evidence for intentional signalling and the potential to learn and generate novel gestures (Corballis 2002; Arbib et al. 2008; Tomasello 2008). Gestural theories often emphasize the *absence* of these language-like qualities in primate vocal or facial communication in order to highlight weaknesses in 'opposing' theories. In this vein, vocalizations and facial expressions, as opposed to gestures, are often described as inflexible, unintentional and involuntary expressions of internal affective states (Tomasello 2008). Crucially, the limited evidence for the production or learning of novel calls or facial expressions is argued to indicate constrained and closed repertoires.

#### *Vocal Theories of Language Evolution*

Theories supporting a vocal origin of language suggest that language built directly on the vocal abilities of our ancestors and that there are commonalities in primate vocalizations and human language (Zuberbühler 2005). Vocal theories also rely on comparative evidence of primate communication. These theories highlight that primate calls can refer to external objects and events (Seyfarth et al. 1980) possibly representing a precursor to human referential abilities. Recent research has revealed that vocalizations can be combined into sequences that are produced in response to specific contexts (Ouattara et al. 2009a), indicating that simple rule-based combinations may exist in primate vocal behaviour (Ouattara et al. 2009b). It is argued that primate vocalizations can be perceived as discrete signals (Fischer 1998), and thus have the potential to be combinatorial, a key feature of language. Vocal theories emphasize that primate call perception is complex and the ability to assign meaning to calls and call combinations is highly flexible (Cheney & Seyfarth 2005; Arnold & Zuberbühler 2008). It is thus in terms of call perception rather than call production that the greatest

similarities with human language are apparent (Cheney & Seyfarth 2005). The absence of these features in primate gestural communication is used to weaken gestural theories in comparison to vocal theories.

#### *Facial Theories of Language Evolution*

Very few theories examine the role that facial expressions might have played in the evolution of human language and those that do treat facial expressions more as a component of gesture or vocalization rather than a different modality. For example, mirror neurons are responsive to mouth-communicative movements in monkeys as well as manual actions and these neurons have been implicated in language evolution (Fogassi & Ferrari 2007). Recent evidence has also shown facial expressions to be perceived by monkeys in an integrated fashion with vocalizations (Ghazanfar & Logothetis 2003). When examined alone facial expressions are usually assumed to be tightly linked to the expression of basic emotions (Ekman 1992), and indeed the *expression* of emotional state is often an implicit component of common definitions of facial expression (e.g. Darwin 1872; Chevalier-Skolnikoff 1973; Ekman 1994). However, Ekman & Friesen (1978) saw the dangers of relying on (assumed) emotional correlates when describing the physical features of facial expression, and so developed the now heavily used, standardized coding system based on muscle movements (The Facial Action Coding System: FACS). The connection between emotion and facial expression has probably influenced the suggestion that facial expressions are largely innate and involuntarily produced, and therefore unimportant for understanding language evolution (e.g. Tomasello 2008). In some ways this might reflect a legacy left by Darwin (1872), who adopted a conversely 'anti-Darwinian' approach by suggesting that facial expressions are not communicative adaptations, but inevitable by-products of internal, emotional states (Fridlund 1992). Only through the introduction of a behavioural ecological approach did the communicative nature of facial expression start to sit within an adaptationist paradigm, and has sometimes been considered a 'para-language' (e.g. Fridlund 1991).

### **EMPIRICAL RESEARCH ON PRIMATE COMMUNICATION**

It is clear that all of these theories of language evolution rely on comparative evidence of primate communication as a central and essential line of argument. This essay is not aimed at evaluating the theories, but instead at examining the distribution of the comparative evidence on which they rely, so a more complete picture can be developed. We also aim to evaluate whether the aforementioned comparisons and claims can legitimately be made given the current body of evidence. We hypothesized that because of different methodologies, availability of technology and theoretical assumptions, there would be several important differences in the type of research conducted in each modality and that combining data, ideas and theories from different modalities might yield a better understanding than each can provide alone.

To gain an accurate picture of the current state of primate communication and to identify where biases in our knowledge may lie, we conducted a systematic search for research conducted on spontaneous primate communication (i.e. we did not include studies in which primates were asked to interpret human signals or use artificial language systems, as these studies make a different contribution to the language evolution debate than studies of spontaneous communication) from 1960 to 2008, using three major relevant search engines (Web of Science, ScienceDirect and PrimateLit; search terms were: 'facial communication' OR 'facial expression\*' OR 'facial display\*' OR 'display AND behaviour' OR 'gestur\*' OR 'gestur\* communication' OR 'gestur\* display\*' OR 'vocalization\*' OR

'vocalization\*' OR 'call' OR 'vocal communication' OR 'vocal\*' AND primate\* OR ape\* OR monkey\* OR macaque\* OR gorilla\* OR orang-utan\* OR baboon\* OR vervet\* OR chimpanzee\* OR gibbon\*). Only empirical, peer-reviewed articles published in English were considered. This provided a cross-section of empirical research completed across modalities (although as keywords can vary considerably it is unlikely to be exhaustive). Each study returned by these search terms was first assessed in terms of the primary modality of communication investigated (vocal, gestural, facial or multimodal). We only included studies that examined one or more of the three main modalities of interest (vocal, gestural and facial communication): studies assessing olfactory, chemical or tactile communication were excluded, with the exception of one study that examined vocal and olfactory communication in an integrated multimodal manner. Although communication in both gestural and facial modalities uses the visual channel, they have been assessed separately by language evolution theories, and so we also examined them separately. We classified a study as multimodal if it assessed relationships or interaction between communicative signals from different modalities in an integrated fashion (see Table 1). For each study we also examined whether the research was observational or experimental, the species class studied (great ape, lesser ape, monkey or prosimian), the research environment (wild or captive) and the focus of the research (producer or receiver).

We examined a total of 553 studies, the vast majority of which were concerned with only a single modality: only 5% of studies examined multimodal communication in an integrated way (see Table 1). The following analyses focus on the remaining unimodal studies. There was a great imbalance in the numbers of studies conducted across modalities with vocal research accounting for 64% of studies, facial 22% and gestural only 9%. As Table 2 and Fig. 1 show, there were significant differences in the species class most commonly used across the three modalities, with vocal research heavily focused on monkey species and, at the other extreme, gestural research heavily focused on great ape species. The relative balance of observational and experimental work varied significantly across modalities ( $\chi^2_2 = 25.80$ ,  $P < 0.0001$ ) with vocal studies being the most experimental (62%) and facial expressions the least (36%).

Table 2 illustrates that most of the primate communication work was conducted in captivity. The research conducted in the wild was distributed unequally across modalities with significantly more than expected focusing on vocal communication (89%;  $\chi^2_2 = 55.26$ ,  $P < 0.0001$ ). Studies using experimental methods in the wild focused solely on vocalizations. In addition, a significantly higher proportion of the facial (39.3%) and vocal (37.5%) studies considered the role of the receiver, compared to gestural studies (19.6%;  $\chi^2_2 = 6.86$ ,  $P = 0.032$ ). Importantly, 100% of gestural studies examined behaviour in terms of the producer, indicating that their central focus was signal production.

In summary, data on gestures were commonly generated from captive apes, using both experimental and observational methods, but with a specific focus on the behaviour of the producer. Facial expression data were mostly observational and generated from captive monkeys but with a focus on both receivers and producers.

**Table 1**

The combined modalities examined by the 28 multimodal studies found by the literature search

Modalities examined in an integrated fashion	Number of studies
Facial, vocal and gestural	9
Facial and gestural	4
Facial and vocal	9
Vocal and gestural	5
Vocal and olfactory	1

**Table 2**

Percentage of the studies within in each modality that were conducted in different research environments, using different approaches, foci and species classes

		Vocal (N=352)	Gestural (N=51)	Facial (N=122)	Integrated multimodal (N=28)
Research environment	Wild	38.4	7.8	8.2	10.7
	Captive	62.8	92.2	95.1	89.3
Approach	Observational	46.3	52.9	65.6	50.0
	Experimental	62.2	49.0	36.0	50.0
Species class*	Great apes	8.5	78.4	23.8	39.3
	Lesser apes	4.5	2.0	3.3	3.6
	Monkeys	83.2	19.6	77.9	53.6
	Prosimians	4.3	0	4.9	3.6
Research focus	Producer	73.9	100	88.5	67.9
	Receiver	37.5	19.6	39.3	46.4

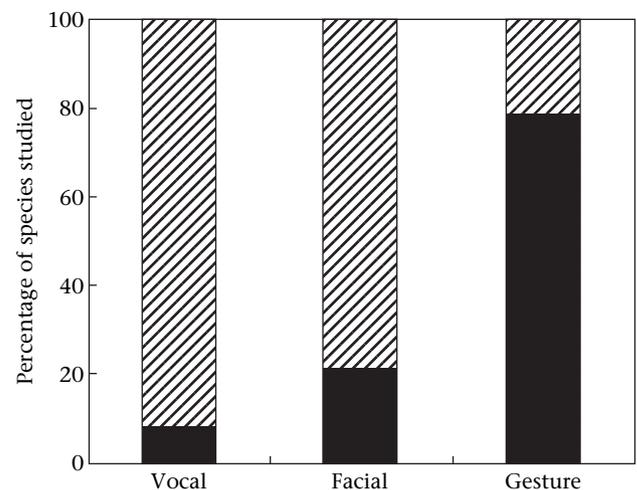
Single studies may have used multiple research environments, approaches, foci and species classes; therefore the sum of these percentages sometimes exceeds 100%.

\* Absolute variation across the categories of this variable will have been influenced by the search terms used: in addition to primate, monkey and ape, we used chimpanzee, orang-utan, gorilla, gibbon, macaque, vervet and baboon in the search terms. The low absolute number of prosimian studies in particular is probably partly due to the absence of lemur or prosimian in the search terms. This should not, however, have affected the relative distribution of these species across modalities.

Vocalization data were mostly generated from monkeys, used experimental methods, and had a broader focus on wild and captive populations, as well as both producer and receiver. Thus, there was a relative lack of vocal research on apes, gestural work conducted in the wild and facial research conducted of an experimental nature. There were also fewer studies that focused on receiver behaviour and responses, compared to producer behaviour, across all modalities.

#### Empirical Evidence and Theories of Language Evolution

The differences in the methodological approaches and research settings used across modalities will inevitably reveal some characteristics of communication in one modality and not the other. For instance, vocal work has historically focused on communication in evolutionary urgent contexts in the wild, such as predator defence, whereas gesture work has focused more on signals in relaxed social contexts in captivity, such as play. There may be a greater



**Figure 1.** Percentage of great ape (black) and other species (hatched; monkeys, prosimians and lesser apes) studied in vocal ( $N = 354$ ), facial ( $N = 134$ ) and gestural ( $N = 51$ ) modalities. Each study could focus on multiple species classes; thus the percentage is in terms of the total number of different species classes studied in each modality. A chi-square test revealed a significant difference in the use of great ape species across the three modalities ( $\chi^2_2 = 144.79$ ,  $P < 0.0001$ ).

evolutionary pressure for unambiguous signals in urgent contexts, such as predator defence, rather than nonurgent social contexts (Tomasello & Zuberbühler 2002). Flexibility in communicative signal production (e.g. producing the same signal in a variety of contexts) may occur only in these nonurgent contexts, in which the receiver can use contextual information to respond in the most adaptive way to the signal (Smith 1977). Currently, gesture work is therefore more likely to reveal flexible signal production than vocal work. In addition the relative absence of experimental work in gestural and facial signals is probably due to current methodological constraints, meaning it is difficult to examine conspecific gestures or facial expressions experimentally. It is easier, for instance, to present previously recorded signals to group members to assess receiver understanding in the vocal modality than it is in the facial or gestural modality.

In some cases, commonly used methodological approaches may also reflect a priori assumptions about the modality, which may itself add an element of circularity to the research by influencing findings, and as such needs to be addressed. For instance, the availability of playback methodologies has provided a large body of evidence indicating that vocalizations function referentially (Fedurek & Slocombe, *in press*). Many vocal researchers therefore look for context-specific call production and, if they fail to find it, the investigation may simply not be published, rather than being presented as evidence for flexible use of a communicative signal across contexts, which is often reported for production of gestures (Pollick & de Waal 2007). Vocal research, in particular, tends therefore to overlook Smith's (1977) argument that displays are dependent on contextual sources of information. Facial expression research is also constrained by the assumption that facial signals are emotional, as this may direct researchers away from cognitive questions.

It is clear there are numerous contributing factors to the gaps in our understanding of primate communication. Until we have a more complete picture of primate communication across modalities and more comparable research results, it is not possible to state unequivocally that gestures are more or less intentional than vocalizations, or that facial expressions are more or less emotional than gestures, and so on. It is, however, essential that absence of evidence for a trait in a poorly researched area is not automatically interpreted as an absence of ability. This is particularly important when absence of a trait is taken to support or refute a particular theory of language evolution. Yet, many theories of language evolution currently hinge on such arguments, and use them to propose which modality was a more likely candidate as a language precursor. We propose that given the considerable holes in our current knowledge of primate communication, such comparisons and theories should be viewed with caution.

## THE WAY FORWARD

We propose several ways in which empirical comparative research into primate communication could help resolve the problems highlighted above and therefore form a more solid base from which theories of language evolution can grow. These proposals add to a growing number of suggestions concerning how to advance the field of language evolution productively, including those advocating the production of precise testable models (e.g. Brighton et al. 2006) and hypotheses (e.g. Fitch 2005).

Unimodal research will continue to be the only option in some contexts partly because of current methodological constraints. This research will be particularly valuable if it first adopts approaches that are as similar as possible across modalities (e.g. examining the same species, in similar research environments, using the same focus of producer/receiver behaviour, etc.), to allow more valid cross-modal comparisons. In addition, the use of the same methods

to compare modalities directly would be very beneficial. For example, match-to-sample touch screen methods have been used successfully with several species, and can be employed to examine both visual and auditory signals (Parr 2004). Second, unimodal research could be more beneficial if targeted in an area that is currently understudied (e.g. vocal research in apes, gestural work in the wild, experimental facial work, research on the receiver), as this will allow gaps in our knowledge to be filled in and may even expose some current assumptions about primate communication to be false. It is also possible that there are other aspects of communicative behaviour that have not been reviewed here and yet also remain understudied; the role of body posture, for example (but see de Gelder & Partan 2009).

Research that addresses the behavioural response of the receiver will continue to be particularly important, as it has long been acknowledged in wider animal communication research that signals are largely shaped through 'receiver psychology' (Guilford & Dawkins 1991). The perceptual system of the receiver acts as an important selection pressure in the evolution of signal structure, with signals the receiver can easily detect, discriminate and remember being most likely to be selected (Rowe 1999). It is also important to note that receivers usually make sense of the signals they receive in context, and often the 'meaning' a receiver extracts from a signal is dependent on both the signal and the context (Smith 1977). Contextual information includes previous social interactions (e.g. 'historical context' Smith 1965) and such social knowledge is likely to be obtained through a variety of modalities (e.g. Cheney & Seyfarth 2005). In this sense, studies examining unimodal signals embedded in the social context could be seen as multimodal; however, more explicit testing of the nature of cognitive integration performed by these primates would be beneficial.

## Multimodal Research

We also suggest that, where possible, researchers should pursue a more integrated, multimodal approach to primate communication, especially where established methods are available. Multimodal signals have been well documented in a number of nonprimate species. They are produced by a variety of species in courtship displays (e.g. wolf spiders, *Schizocosa ocreata*: Uetz et al. 2009), agonistic displays (e.g. frogs, *Allobates femoralis*: de Luna et al. 2010) and some birds also combine visual signals with song or call production (e.g. Beer 1976; O'Loughlen & Rothstein 2010). Insects commonly combine a visual coloration warning signal with an olfactory or auditory signal to produce a multimodal warning signal to predators and research shows that the combination of signals increases avoidance by insect-eating birds (Rowe & Guilford 1996). Similarly, in Californian ground squirrels, *Otospermophilus beecheyi*, the integration of visual and vocal signals plays a crucial role in communication about predators (Owings & Virginia 1978; Hennessy et al. 1981). Recent work presenting robotic models to wild tree squirrels, *Sciurus carolinensis*, has established that although unimodal visual or auditory signals do elicit antipredator behaviour in receivers, the strongest responses come from the combined multimodal signal (Partan et al. 2009). It is not just that multimodal signals often elicit the strongest response from receivers, Rowe (1999) also reported that receiver responses to multimodal signals can be synergistic in that they are often stronger than the summation of the responses to individual components; thus we are unlikely to understand multimodal signals fully by studying the component signals independently. This literature highlights the importance of interactions across signalling modalities for successful communication and this message is echoed by the multimodal studies that have been conducted with primates.

In primates, match-to-sample paradigms have revealed that chimpanzees, *Pan troglodytes*, can recognize combined vocal and facial signals cross-modally, but one modality is more salient depending on the specific social function of the display (Parr 2004). For example, long-distance pant-hoot calls are more readily identified in the auditory modality, and screams visually. From a production perspective, chimpanzees also produce visual, tactile and auditory signals to communicate with human caretakers and will flexibly choose or combine these different types of signals as a function of the attentional state of the human (Leavens et al. 2010). In her comprehensive analysis of multichannel signal composition in rhesus macaques, *Macaca mulatta*, Partan (2002) found that some cross-modal behavioural components are almost always associated (screams with prototypical grimace), and others are more flexibly combined (barks with ears retracted). Flexible associations suggest greater potential for intentional signalling, and thus are highly interesting from a language perspective. At a neuronal level, we know that various brain regions, including the auditory cortex and the superior temporal sulcus, contain neurons that respond to the integration of visual and auditory signals, such as dynamic vocalizing monkey faces (Ghazanfar & Schroeder 2006). Multimodal research, as illustrated by the above exemplars, is vital as it demonstrates that primate communication is inherently multimodal, at both a behavioural and neuronal level, indicating that unimodal research tells only part of the complex story.

Research on human communication also strongly suggests that language is exchanged in a multimodal format. Humans constantly augment language with nonlinguistic and often nonverbal cues (through face, hands and body), which can fundamentally adjust the message conveyed. For instance, eyebrow movements punctuate speech for emphasis (Ekman 1979) and gestures automatically accompany speech in the absence of learning (Iverson & Goldin-Meadow 1998). Rather than representing emotional vestiges that need to be stripped from language in order to expose the fundamental cognitive components, these nonverbal signals are part of an important composite message.

It is clear that both human language and primate communication are intrinsically multimodal and in searching for unimodal precursors to language, scientists are overlooking the complexity that is inherent in this multimodal system (Partan & Marler 1999). Abandoning the traditional distinctions between gesture, facial expression and vocalization could therefore have a large and positive impact on the study of language evolution. First, by studying modalities side by side or examining the contribution of each modality to successful communication, scientists will use comparable methods and contexts for each modality, which is vital for the generation of data for the purpose of comparison. Such multimodal data will not be subject to the biases we have highlighted in the current unimodal data and can therefore be used by theorists to make valid comparisons across modalities. Data from multimodal studies may reveal a different pattern of results to unimodal research, on which most theories currently stand and this may lead to modification of these theories.

Second, the small but important amount of multimodal research conducted to date highlights the continuity in multimodal communication across human and primate species. This suggests that language may have evolved through an integrated combination of vocal, gestural and facial communication, rather than a unimodal system. If so, an acknowledgment of this could lead to the development of new theories of language evolution or provide further support for existing theories. For example, Dunbar (1993) proposed that language evolved to replace social grooming (which bonds social groups in primates), and was more efficient in this role as it allowed humans to produce the reinforcing, social-bonding effects of grooming (through opiate production), but at a distance. Dunbar suggested that language achieves this not only through information

transfer and gossip, but also through emotional means (e.g. laughter, facial expression). It is possible, therefore, that communicating simultaneously through a range of modalities is the skill that truly occupies the functional niche of primate grooming, and not the cognitive aspect alone. When considering the precursors to language, theories that ignore the components deemed to be less cognitive could therefore be missing a vital component.

In short, further multimodal research could support the development of new theories of language evolution that avoid the traditional gestural/vocal/facial divide and instead focus on the evolution of a multimodal communication system.

## Acknowledgments

We thank the University of Portsmouth, Department of Psychology Research Committee, the University of York Pump-priming fund and the BBSRC for funding and the anonymous referees for their very helpful comments on the manuscript.

## References

- Arbib, M. A., Liebal, K. & Pika, S. 2008. Primate vocalization, gesture, and the evolution of human language. *Current Anthropology*, **59**, 1053–1076.
- Arnold, K. & Zuberbühler, K. 2008. Meaningful call combinations in a non-human primate. *Current Biology*, **18**, R202–R203.
- Beer, C. 1976. Some complexities in the communication behavior of gulls. *Annals of the New York Academy of Sciences*, **280**, 413–432.
- Bickerton, D. 1992. *Language and Species*. Chicago: University of Chicago Press.
- Brighton, H., Mata, R. & Wilke, A. 2006. Reconciling vague and formal models of language evolution. *Behavioral and Brain Sciences*, **29**, 282.
- Cheney, D. L. & Seyfarth, R. M. 2005. Constraints and preadaptations in the earliest stages of language evolution. *Linguistic Review*, **22**, 135–159.
- Chevalier-Skolnikoff, S. 1973. Facial expression of emotion in nonhuman primates. In: *Darwin and Facial Expressions* (Ed. by P. Ekman), pp. 11–89. New York: Academic Press.
- Corballis, M. C. 2002. *From Hand to Mouth, the Origins of Language*. Princeton, New Jersey: Princeton University Press.
- Darwin, C. 1872. *The Expression of the Emotions in Man and Animals*. 3rd edn. New York: Oxford University Press.
- Dunbar, R. I. M. 1993. Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, **16**, 681–735.
- Ekman, P. 1979. About brows: emotional and conversational signals. In: *Human Ethology: Claims and Limits of a New Discipline* (Ed. by M. von Cranach, F. Koppa, W. Lepenies & D. Ploog), pp. 169–202. New York: Cambridge University Press.
- Ekman, P. 1992. An argument for basic emotions. *Cognition and Emotion*, **6**, 169–200.
- Ekman, P. 1994. Moods, emotions and traits. In: *The Nature of Emotion: Fundamental Questions* (Ed. by P. Ekman & R. J. Davidson), pp. 56–58. Oxford: Oxford University Press.
- Ekman, P. & Friesen, W. V. 1978. *The Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, California: Consulting Psychologists Press.
- Endress, A. D., Cahill, D., Block, S., Watumull, J. & Hauser, M. D. 2009. Evidence of an evolutionary precursor to human language affixation in a non-human primate. *Biology Letters*, **5**, 749–751.
- Fedurek, P. & Slocombe, K. E. In press. Primate vocal communication: a useful tool for understanding human speech and language evolution? *Human Biology*.
- Fischer, J. 1998. Barbary macaques categorize shrill barks into two call types. *Animal Behaviour*, **55**, 799–807.
- Fitch, W. T. 2005. The evolution of language: a comparative review. *Biology and Philosophy*, **20**, 193–230.
- Fogassi, L. & Ferrari, P. F. 2007. Mirror neurons and the evolution of embodied language. *Current Directions in Psychological Science*, **16**, 136–141.
- Fridlund, A. J. 1991. Evolution and facial action in reflex, social motive, and para-language. *Biological Psychology*, **32**, 3–100.
- Fridlund, A. J. 1992. Darwin's anti-Darwinism in the expression of the emotions in man and animals. In: *International Review of Studies on Emotion* (Ed. by K. T. Strongman), pp. 117–137. New York: J. Wiley.
- Gardner, R. A. & Gardner, B. T. 1969. Teaching sign language to a chimpanzee. *Science*, **165**, 664–672.
- de Gelder, B. & Partan, S. 2009. The neural basis of perceiving emotional bodily expressions in monkeys. *Neuroreport*, **20**, 642–646.
- Ghazanfar, A. A. & Logothetis, N. K. 2003. Neuroperception: facial expressions linked to monkey calls. *Nature*, **423**, 937–938.
- Ghazanfar, A. A. & Schroeder, C. E. 2006. Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, **10**, 278–285.
- Guilford, T. C. & Dawkins, M. S. 1991. Receiver psychology and the design of animal signals. *Animal Behaviour*, **42**, 1–14.

- Hauser, M., Chomsky, N. & Fitch, W. T.** 2002. The faculty of language: what is it, who has it, and how did it evolve? *Science*, **298**, 1569–1579.
- Hayes, C.** 1951. *The Ape in our House*. New York: Harper & Brothers.
- Hennessy, D. F., Owings, D. H., Rowe, M. P., Coss, R. G. & Leger, D. W.** 1981. The information afforded by a variable signal: constraints on snake-elicited tail flagging by California ground squirrels. *Behaviour*, **78**, 188–226.
- Hewes, G. W.** 1973. Primate communication and the gestural origin of language. *Current Anthropology*, **12**, 5–24.
- Hockett, C. F.** 1960. The origin of speech. *Scientific American*, **203**, 89–97.
- Iverson, J. M. & Goldin-Meadow, S.** 1998. Why people gesture as they speak. *Nature*, **396**, 228.
- Lachlan, R. F. & Feldman, M. W.** 2003. Evolution of cultural communication systems: the coevolution of cultural signals and genes encoding learning preferences. *Journal of Evolutionary Biology*, **16**, 1084–1095.
- Leavens, D. A., Russell, J. L. & Hopkins, W. D.** 2010. Multimodal communication by captive chimpanzees (*Pan troglodytes*). *Animal Cognition*, **13**, 33–40.
- de Luna, A. G., Hodl, W. & Amezcua, A.** 2010. Colour, size and movement as visual subcomponents in multimodal communication by the frog *Allobates femoralis*. *Animal Behaviour*, **79**, 739–745.
- O’Loghlen, A. L. & Rothstein, S. I.** 2010. Multimodal signalling in a songbird: male audiovisual displays vary significantly by social context in brown-headed cowbirds. *Animal Behaviour*, **79**, 1285–1292.
- Ouattara, K., Lemasson, A. & Zuberbühler, K.** 2009a. Campbell’s monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences, U.S.A.*, **106**, 22026–22031.
- Ouattara, K., Lemasson, A. & Zuberbühler, K.** 2009b. Campbell’s monkeys use affixation to alter call meaning. *PLoS One*, **4**, e7808.
- Owings, D. H. & Virginia, R. A.** 1978. Alarm calls of California ground squirrels (*Spermophilus beecheyi*). *Zeitschrift für Tierpsychologie*, **46**, 58–70.
- Parr, L. A.** 2004. Perceptual biases for multimodal cues in chimpanzee (*Pan troglodytes*) affect recognition. *Animal Cognition*, **7**, 171–178.
- Partan, S. R.** 2002. Single and multichannel signal composition: facial expressions and vocalizations of rhesus macaques (*Macaca mulatta*). *Behaviour*, **139**, 993–1027.
- Partan, S. R. & Marler, P.** 1999. Communication goes multimodal. *Science*, **283**, 1272–1273.
- Partan, S. R., Larco, C. P. & Owens, M. J.** 2009. Wild tree squirrels respond with multisensory enhancement to conspecific robot alarm behaviour. *Animal Behaviour*, **77**, 1127–1135.
- Patterson, F. G.** 1978. The gestures of a gorilla: language acquisition in another pongid. *Brain and Language*, **5**, 72–97.
- Pollick, A. S. & de Waal, F. B. M.** 2007. Ape gestures and language evolution. *Proceedings of the National Academy of Sciences, U.S.A.*, **104**, 8184–8189.
- Rizzolatti, G. & Arbib, M. A.** 1998. Language within our grasp. *Trends in Neurosciences*, **21**, 188–194.
- Rowe, C.** 1999. Receiver psychology and the evolution of multicomponent signals. *Animal Behaviour*, **58**, 921–931.
- Rowe, C. L. & Guilford, T.** 1996. Hidden colour aversions of domestic chicks triggered by pyrazine odours of insect warning displays. *Nature*, **383**, 520–522.
- Seyfarth, R. M., Cheney, D. L. & Marler, P.** 1980. Monkey responses to three different alarm calls: evidence for predator classification and semantic communication. *Science*, **210**, 801–803.
- Smith, W. J.** 1965. Message, meaning and context in ethology. *American Naturalist*, **99**, 405–409.
- Smith, W. J.** 1977. *The Behavior of Communicating: An Ethological Approach*. Cambridge, Massachusetts: Harvard University Press.
- Tomasello, M.** 2008. *Origins of Human Communication*. Cambridge, Massachusetts: MIT Press.
- Tomasello, M. & Zuberbühler, K.** 2002. Primate vocal and gestural communication. In: *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition* (Ed. by M. Bekoff, C. S. Allen & G. Burghardt), pp. 293–299. Cambridge, Massachusetts: MIT Press.
- Uetz, G. W., Roberts, J. A. & Taylor, P. W.** 2009. Multimodal communication and mate choice in wolf spiders: female response to multimodal versus unimodal signals. *Animal Behaviour*, **78**, 299–305.
- Zuberbühler, K.** 2005. The phylogenetic roots of language: evidence from primate communication and cognition. *Current Directions in Psychological Science*, **14**, 126–130.