Neuroscience and Biobehavioral Reviews 80 (2017) 201–209

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review article

Taking turns across channels: Conversation-analytic tools in animal communication

Marlen Fröhlich[a,b,*]

[a] Humboldt Research Group 'Evolution of Communication', Max Planck Institute for Ornithology, 82319 Seewiesen, Germany
[b] Anthropological Institute and Museum, University of Zurich, 8057 Zurich, Switzerland

ARTICLE INFO

Keywords:
Animal communication
Conversation analysis
Language origins
Multimodal signaling
Turn-taking

ABSTRACT

In the quest to bridge the gulf between the fields of linguistics and animal communication, interest has recently been drawn to turn-taking behavior in social interaction. Vocal turn-taking is the core form of language usage in humans, and has been examined in numerous species of birds and primates. Recent studies on great apes have shown that they engage in a bodily form, gestural turn-taking, to achieve mutual communicative goals. However, most studies on turn-taking neglected the fact that signals are prevalently perceived and produced in a multimodal format. Here, I propose that research on animal communication may benefit a more holistic and dynamic approach: studying turn-taking using a multimodal, conservation-analytic framework. I will discuss recent comparative research that implemented this approach via a specific set of parameters. In sum, I argue that a conversation-analytic framework might help substantially to pinpoint the ways in which crucial components of language are embodied in the 'human interaction engine'.

1. How to bridge the gulf between human and nonhuman communication?

1.1. Language is more than speech

Humans communicate in an utterly unique way: We use language as an abstract symbolic code to share information directly (Tomasello, 2008). It has been suggested that it is especially the cooperative nature of our communication system, or the capacity to mutually donate information, which sets us apart from the rest of the animal kingdom (Tomasello, 2008; van Schaik, 2016). While commonly narrowed down to speech, language is in fact an intrinsically multimodal phenomenon. Spoken language is tightly interlinked with visual signals and cues across cultures, ages and tasks (Levinson and Holler, 2014; Iverson and Goldin-Meadow, 1998; McNeill, 2000). Speech-accompanying manual gestures, body postures, facial expressions and gaze direction can provide salient information about the signaler's identity, emotional state and intentions, as well as social hierarchy (Tomasello and Camaioni, 1997). McGurk and MacDonald (1976) demonstrated that the perception of speech sounds by adults is modulated by the observation of the accompanying lip movements (termed 'McGurk effect'). Moreover, Massaro's perceptual experiments (Massaro, 1998; Massaro and Egan, 1996), manipulating the degree of conflict in audio and visual speech information, suggested that humans rely on both channels to understand the signal, giving more weight to the channel with the most reliable information. Nonverbal cues can fundamentally adjust the conveyed message, for instance when eyebrow movements punctuate speech for emphasis (Ekman, 1979). It has also been shown that humans' speech-accompanying ('co-speech') manual gestures possess a crucial communicative function in resolving ambiguity in spoken language (Holle and Gunter, 2007), and add a substantial amount of information to the speaker's message (e.g., Graham and Argyle, 1975; Goldin-Meadow, 1999; Kelly et al., 1999).

1.2. The comparative approach

While language is undoubtedly unique with regard to the complexity of underlying semantic and syntactic structures, it clearly has a strong biological foundation evident in the brain and the vocal apparatus, the crucial elements for speech production. Understanding language evolution has been hampered by the fact that these anatomical features do not fossilize (Ghazanfar and Rendall, 2008). Language as a whole is unique to humans, but there is ample evidence that many critical components of language are shared with other animals (Liebal et al., 2013; Hauser et al., 2002; Levinson and Holler, 2014). When searching for the evolutionary origins of human communication it is thus imperative to understand the selection pressures acting on and the cognitive abilities underlying the communicative behavior in nonhu-

* Corresponding author at: University of Zurich, Anthropological Institute and Museum, Winterthurerstrasse 190, 8057 Zurich, Switzerland.
E-mail address: marlen.froehlich@uzh.ch.

http://dx.doi.org/10.1016/j.neubiorev.2017.05.005
Received 11 January 2017; Received in revised form 21 March 2017; Accepted 8 May 2017
Available online 10 May 2017

man species (Hauser et al., 2002). In the past decades the comparative approach has been widely used to gain insight into the cognitive building blocks and selective pressures that shaped the human communication system (e.g., Arbib et al., 2008; Tomasello, 2008; Hauser et al., 2002). Crucial ingredients of our linguistic communication system and established measures of 'communicative complexity', such as intentionality, referentiality and compositional syntax, have been demonstrated for a wide range of taxa of nonhuman species, (e.g., Vail et al., 2013; Engesser et al., 2016; Pollard and Blumstein, 2012). Nonetheless, the majority of comparative work stems from humans' closest living relatives, the great apes (e.g., Call and Tomasello, 2007; Pika and Liebal, 2012; Liebal et al., 2013; Arbib et al., 2008). The two *Pan* species—chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*)—have been most commonly used as models for the origins of human behavior. This research bias has been justified by the relatively recent divergence of the human from the *Pan* lineage around five to eight million years ago (Prufer et al., 2012; Langergraber et al., 2012), and the complex social relationships found in these sister species that are navigated by means of a large signal diversity. However, the extent to which the components of language are built upon socio-cognitive skills in other animals are a matter of fierce debate (Scott-Phillips, 2015; Townsend et al., 2016; Fischer and Price, 2016; Moore, 2015; King, 1999). Research into language precursors has been hampered by the lack of agreement in approaches between human and nonhuman communication research, but also between vocal and gestural research (Liebal et al., 2013; Townsend et al., 2016; Slocombe et al., 2011).

1.3. Enriching the information-transmission paradigm

For many decades, Shannon's information-transmission metaphor (Shannon, 1948) has dominated our view of communication: the information source or sender produces a signal via a transmitter. This signal is transmitted through a single channel, which is impacted by a noise source, before it is reaches the destination or receiver via sensory receptors (see also Fig. 1). While such a visualisation has been undoubtedly useful to exchange ideas and improve understanding, it dismisses numerous complexities of communication since in real communicative interactions *in situ*, a wealth of factors might play a role (Partan, 2013). In addition to the multiple components of the signal itself and the physical environment, each interaction is strongly affected by the social environment (i.e. bystanders), the social context, the relationship between the interactants as well as prior interactional experiences (familiarity) (Fröhlich et al., 2017, 2016b; Partan, 2013). Even if we study communication only at a dyadic level, it is in many cases not a straightforward task to determine which of the interactants initially represented the signaler and which the receiver, especially if the behavior of interest is produced synchronically and in a dynamic and cooperative exchange (Fröhlich et al., 2016a; Partan, 2013).

In line with this premise, qualitative studies proposed that signaling is not only an inherently multimodal but also highly dynamic process that is strongly influenced by contextual factors (King and Shanker, 2003). Shanker and King (King and Shanker, 2003; Shanker and King, 2002) claimed that the information-transmission framework, which treats communication as a *sequential* process in which partners take turns in signaling (Argyle, 1988), limits our understanding of the dynamic, co-regulated interactions that occur during communication. However, an informational approach is still regarded suitable to examine complex signaling systems in fine detail (Partan, 2013; Zentall, 2002) and could be principally applied to multiple aspects of the same interactive and dynamic system. While in many cases it seems appropriate to consider communicative interactions as 'dances' (Shanker and King, 2002), the reductions of uncertainty must be identified in any state of an information-processing network that comprises the entire system. The empirical challenge of identifying signaler and receivers remains one of the essential issues in identifying the factors contributing to the communicative nature of any interaction.

2. A multimodal approach in animal communication

2.1. Multimodal research in non-primate and primate species

Animals communicate with con- and hetero-specifics using multicomponent signals in every sensory modality, employing visual, audible, olfactory and other signals (Higham and Hebets, 2013; Partan, 2013). Multimodal signaling has been studied intensively in a large number of non-primate species (for reviews see Partan and Marler, 2005; Hebets and Papaj, 2005; Guilford and Dawkins, 1991; Billen, 2006; Brumm and Slabbekoorn, 2005; Starnberger et al., 2014), and here I will introduce the topic by using the example of courtship signals in different taxa. Wolf spiders integrate complex visual and seismic signals in their courtship displays, with females detecting males faster if they use multimodal rather than unimodal signals (Uetz et al., 2009). Female túngara frogs show a preference for multimodal courtship displays, consisting of the vocalization and the visual signal of the inflating air sac, compared to unimodal displays (Taylor et al., 2008). In line with these findings, it was shown that female pigeons are more responsive to a multisensory audio/video courtship signal than to either component alone (Partan et al., 2005). Contrarily, Smith and Evans (2008, 2009) found that the visual and auditory components of the multimodal tidbitting display in male fowls were redundant, that is both the multimodal and unimodal signals elicit the same female response. Thus, multimodal signals are thought to enhance either efficacy or information content (Partan and Marler, 1999; Hebets and Papaj, 2005; Guilford and Dawkins, 1991); this will be discussed in more depth below.

Primate communication, however, has rarely been studied holistically, with most studies focusing on either gesture or vocalization alone (Slocombe et al., 2011). A detailed review by Slocombe et al. (2011) revealed that until 2011 only about five per cent of primate communication studies focused on more than one modality. However, neither apes nor humans gesture or vocalize in isolation; in fact, research on primate communication at both the behavioral and neuronal level has demonstrated that it is inherently multimodal (for review see Liebal et al., 2013). Cross-modal integration of vocal and visual signals has been shown in macaques (Ghazanfar and Logothetis, 2003) and chimpanzees (Parr, 2004), who were able to match vocalization with the corresponding facial movement that originally accompanied the sound ('audio-visual matching'). This face/voice integration has been investigated in great depth in macaques (e.g., Chandrasekaran et al., 2011; Adachi and Hampton, 2011; Ghazanfar et al., 2005; Habbershon et al., 2013; Sliwa et al., 2011). Moreover, studies in both wild and captive environments revealed that apes often use signal combinations conveying context-specific information that would not be available from a single sensory input (Leavens et al., 2010; Crockford and Boesch, 2003; Genty et al., 2014). Consequently, some researchers suggested that there is continuity in multimodal communication from primates to humans (Gillespie-Lynch et al., 2014; Lameira et al., 2012; Taglialatela et al., 2011).

Unimodal studies are not suitable to unravel the structure and function of complex signals, and thereby the consequences of combining signal components (Liebal et al., 2013). If research efforts focus on single modalities only, we might miss much of the complexity inherent to a communicative system. For instance, the way signals of different modalities are produced, transmitted and received has important consequences for environmental selection pressures, such as sensory conditions, that act on the behavior of animals (Partan and Marler, 1999).

2.2. Towards an inclusive definition of multimodal communication

The definition and operationalization of multimodal signals still lacks consensus among research disciplines. Evolutionary psychologists working on primate communication have mainly focused on the