Cognition 113 (2009) 226–233

Contents lists available at ScienceDirect

Cognition

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

Signalling signalhood and the emergence of communication

Thomas C. Scott-Phillips *, Simon Kirby, Graham R.S. Ritchie

School of Psychology, Philosophy and Language Sciences, University of Edinburgh, Edinburgh EH8 9AD, United Kingdom

ARTICLE INFO

Article history:
Received 21 August 2008
Revised 11 August 2009
Accepted 11 August 2009

Keywords:
Communication
Emergence of communication
Common ground
Language
Evolution
Symbolism
Communicative intent
Dialogue
Embodiment
Embodied communication game

ABSTRACT

A unique hallmark of human language is that it uses signals that are both learnt and symbolic. The emergence of such signals was therefore a defining event in human cognitive evolution, yet very little is known about how such a process occurs. Previous work provides some insights on how meaning can become attached to form, but a more foundational issue is presently unaddressed. How does a signal signal its own signalhood? That is, how do humans even know that communicative behaviour is indeed communicative in nature? We introduce an experimental game that has been designed to tackle this problem. We find that it is commonly resolved with a bootstrapping process, and that this process influences the final form of the communication system. Furthermore, sufficient common ground is observed to be integral to the recognition of signalhood, and the emergence of dialogue is observed to be the key step in the development of a system that can be employed to achieve shared goals.

1. Introduction

Human language is the only communication system in the natural world where the signals are both learnt and symbolic (Deacon, 1997). These twin features give rise to an emergence problem: if there is no relationship between form and meaning, and if meanings are not innately specified, then how can individuals agree on what forms should refer to what meanings in the first place (Oliphant, 2002)? Almost nothing is known about the answer to this question. Previous experimental (de Ruiter, Noordzij, Newman-Norland, Hagoort, & Toni, 2007; Fay, Garrod, MacLeod, Lee, & Oberlander, 2004; Galantucci, 2005; Healey, Swoboda, Umata, & King, 2007; Selten & Warglien, 2007), computational (e.g. Hurford, 1989; Noble, 2000; Nowak & Krakauer, 1999; Smith, 2004) and theoretical studies (e.g. Lewis, 1969) offer some insights; but all have, with one exception (Quinn, 2001), assumed that at the very earliest stages of a system's development individuals are able to detect that a given behaviour is intended to be communicative. Yet this cannot be taken for granted: before potential receivers can access the problem of what a communicative behaviour must mean, they must first recognise that the behaviour is indeed communicative.

The recognition of informative intent is a fundamental component of (non-natural) meaning (Grice, 1971). Yet previous work, whether it is concerned with learnt or innate symbolism, has avoided the question of how this is achieved. This has been done in (at least) one of three ways. First, the communication channel may be pre-defined (e.g. Fay et al., 2004; Galantucci, 2005; Healey et al., 2007). This will evade the issue since participants know that any inputs that come to them via the communication channel are (almost certainly) communicative in nature. Second, the roles of signaller and receiver may be pre-defined (e.g. de Ruiter et al., 2007; Garrod, Fay, Lee, Oberlander, & MacLeod, 2007; Selten & Warglien, 2007). Although this does not make communicative behaviour quite so salient as a pre-defined communication channel,

* Corresponding author. Tel.: +44 0131 650 3956.
E-mail address: thom@ling.ed.ac.uk (T.C. Scott-Phillips).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.08.009

it nevertheless primes the receiver to interpret the signaller's behaviour in communicative terms. Finally, the possible forms that a signal might take may be pre-specified by the researcher, which renders it is immediately recognisable as a signal. Such an approach is inherent in game-theoretic accounts of communication (e.g. Lewis, 1969) but may also be seen in some computational (e.g. Hurford, 1989; Noble, 2000; Nowak & Krakauer, 1999; Smith, 2004) and experimental (e.g. Selten & Warglien, 2007) work. All of these scenarios mean that the problems that are investigated are how to map form onto meanings, and in some cases (e.g. de Ruiter et al., 2007) how to construct forms, but if we wish to study the question of emergence we must address an even more foundational issue: how do (potential) receivers even know that there is a signal? Put another way, *how does a signal signal its own signalhood*? There is one previous study using evolved robots that directly addresses this question (Quinn, 2001), but that work studied the emergence of an innate, iconic system. We, on the other hand, are interested in the emergence of a learnt, symbolic system[1].

If we wish to address this question our investigative set-up must allow communicative behaviour either to emerge from non-communicative behaviour or be created *de novo*. This means, at a minimum, that we must not pre-define the communication channel, the roles of signaller and receiver, or the form space. More generally, the problem's solution must not be an artifact of the experimental design, and we must instead allow communicative behaviour either to emerge from non-communicative behaviour or be created *de novo*. Importantly, therefore, the task should not be one that can be solved with a deductive choice of the most suitable channel from a number of candidate possibilities. Instead, we must insist that participants co-opt their behaviours in the world for communicative purposes. In short, we must demand that communicative behaviour be embodied. In general, to embody is to make concrete or to give physical form to some entity. For cognition, this means that the bodies that are controlled by brains are themselves an integral part of the cognitive process (see Wilson, 2002 for a review of the various ways in which this point may play out). For communication, it means, minimally, that there should be no *a priori* distinction between communicative and non-communicative behaviour. The act of communication must be situated in the world (as that world is defined by the investigative approach). There is at least one previous piece of experimental work with human participants that satisfies this condition (de Ruiter et al., 2007), and that study correspondingly offers insights into the origins of our communicative intentions. However, it is not ideally suited to the present task for two reasons. First, as mentioned above, it pre-defines the roles of signaller and receiver. Second, iconic solutions are possible, and indeed they are found by participants (this is also the case in Galantucci, 2005). Thus in addition to embodiment and the other constraints mentioned above, we also demand that iconicity (and indeed indexicality) be impossible.

This paper introduces the embodied communication game (ECG), an interactive, cooperative two-player game which satisfies these conditions. Pairs of participants must coordinate their behaviour to solve a simple task where they lack shared information, yet they have no interaction with each other except for their movements[2] within the game's world. This means that these movements must perform both tasks necessary to succeed: (i) travelling within the world; and (ii) communication. Consequently, participants must not just agree on what behaviours correspond to what meaning, but when creating these symbols they must find a way to signal that a given behaviour is a signal. For many participants it is not obvious how they can achieve this goal: many of the pairs of participants are unable to find any form of communication whatsoever (see the results section below). This is because the ECG uniquely demands not only that the participants agree on what movements will correspond to what meanings, but that the participants realise that they are able to use their movements to signal to each other at all. Then, once they recognise this, they must find some way to signal the fact that some of their movements are communicative in nature.

2. The embodied communication game

In the ECG each player is represented as a stick man, each located in his own 2×2 box. Each of the four quadrants is coloured either red, blue, green or yellow, at random. Each player sees both boxes, and the movements within them, but can see only the colours of their own box; and both players know that the experience is the same for the other player. At the beginning of each round each players' stick man begins in one of the quadrants of his/her box. This starting point is chosen at random in each round. The players can move between quadrants at will, but each move is from the centre of each quadrant to the centre of the other quadrants, so they are unable to trace out letters or other symbols with their movements. Each press of the arrow buttons takes the stick man directly to the centre of the new quadrant at a fixed speed. The players press the space bar to finish. Once both players have finished the colours of all quadrants are revealed to both players. If they have finished on identically-coloured quadrants they score a point; if not then no point is scored. Both players then press space again and a new round begins. Screenshots of each player's view, both before and after both players have pressed space to finish the round, can be seen in Fig. 1.

The colours of all quadrants are randomly assigned in every round, with the proviso that at least one of the four colours will appear in both boxes, so that it is always in

[1] By *iconic* we mean systems in which the sign bears a resemblance (physical, auditory, etc.) to its referent; *symbolic* systems, in contrast, exhibit arbitrary relationships. As an example of an innate, iconic system we would suggest the aspect of bee dance that refers to the direction of the nectar; and as an example of a learnt, symbolic system we would point to human language.

[2] Our sense of *movement* is actually slightly more broad than just visible movements, and should be construed as 'game moves' which include physical movements and also end-of-turn indicators. These are both embodied in the sense that they are actions required by the player to traverse the space described by the ECG's world.