New Ideas in Psychology 32 (2014) 131–146

Contents lists available at SciVerse ScienceDirect

New Ideas in Psychology

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

Perspective-taking in dialogue as self-organization under social constraints

Nicholas D. Duran*, Rick Dale

Cognitive and Information Sciences, School of Social Sciences, Humanities and Arts, University of California, Merced, CA 95343, United States

Keywords:
Language
Perspective-taking
Dynamical systems
Communication
Spatial cognition

A B S T R A C T

We present a dynamical systems account of how simple social information influences perspective-taking. Our account is motivated by the notion that perspective-taking may obey common dynamic principles with perceptuomotor coupling. We turn to the prominent HKB dynamical model of motor coordination, drawing from basic principles of self-organization to describe how conversational perspective-taking unfolds in a low-dimensional attractor landscape. We begin by simulating experimental data taken from a simple instruction-following task, in which participants have different expectations about their interaction partner. By treating belief states as different values of a control parameter, we show that data generated by a basic dynamical process fits overall egocentric and other-centric response distributions, the time required for participants to enact a response on a trial-by-trial basis, and the action dynamics exhibited in individual trials. We end by discussing the theoretical significance of dynamics in dialog, arguing that high-level coordination such as perspective-taking may obey similar dynamics as perceptuomotor coordination, pointing to common principles of adaptivity and flexibility during dialog.

1. Introduction

Perspectives are a fundamental aspect of daily interaction. In order to be understood when speaking, or in order to understand someone who is speaking, it is often important to integrate attributions we might make about our conversation partner. If someone wanders up and asks in an accent of a non-native speaker, “Where is downtown Merced?” we may adjust how we articulate our instructions, in a way that is shaped by knowledge of this person. When a close friend asks, “How do I look in these pants?” our response may be shaped by knowledge of the person’s traits, their mood on that day, or the gravity of the event to be attended. Sometimes when we violate these principles of perspective-taking, consequences are dire; other times, they can be innocuous.

This process of integrating information about another human being with whom we are talking is one of our most heralded cognitive skills. However, an account of such perspective-taking skill has not been developed in emerging dynamical accounts of interpersonal coordination. Instead, the focus has been on the perceptual and motor channels and how they couple individuals. In experiments motivated by a dynamical systems account of interpersonal interaction, there is clear evidence that people spontaneously coordinate their movements during communicative tasks (Dittmann & Llewellyn, 1969; Fowler, Richardson, Marsh, & Shockley, 2008; Kendon, 1970; Shockley, Richardson, & Dale, 2009). The structure of this coordination indicates that people operate as a coupled system, whereby individual motor systems are reorganized

* Corresponding author. Tel.: +1 901 490 5922.
E-mail addresses: nduran2@ucmerced.edu (N.D. Duran), rdale@ucmerced.edu (R. Dale).

0732-118X/$ – see front matter
http://dx.doi.org/10.1016/j.newideapsych.2013.03.004

into more efficient modes of adaptive responding (Riley, Richardson, Shockley, & Ramenzoni, 2011). These modes are better able to stabilize in the presence of perturbations and in transitioning between shared behavioral repertories. Moreover, in some domains, this alignment has been hypothesized to reflect coupled mental states that facilitate information transmission, as in promoting common frames of reference for language comprehension, and in establishing turn-taking rhythms that assist in word learning (Pereira, Smith, & Yu, 2008; Richardson & Dale, 2005).

Critical to this view of motor and cognitive coordination are the perceptuomotor channels that bind individuals into functional units. Cues such as eye gaze, acoustic patterns of speech, and the movements of another's hands and head, all constrain how behavior systematically unfolds in a social environment (Fowler et al., 2008; Richardson, Marsh, & Schmidt, 2005; Shockley, Santana, & Fowler, 2003). What remains unclear, however, is how the mere beliefs or knowledge about another, rather than observations of their actions, act to coordinate shared cognitive states. In numerous studies, it has been shown that assessments about abstract, other-oriented *attributes*, ranging from another's linguistic efficiency to their geographic region of residence, have immediate effects on how language users come to negotiate and share meaning (see Brennan, Galati, & Kuhlen, 2010 for a review). Put differently, the connection between interacting individuals is not always purely perceptuomotor; sometimes it is *attributional* or *informational*.

In the current work, we seek to bridge what might appear at first glance to be disparate research programs. Dynamical systems accounts typically operate within contexts where language users are physically co-present and bound by similar motor systems. Yet, communication still succeeds when all that links language users is informational content, such as in the extreme case of computer-mediated communication, to everyday conversations where the perceptuomotor cues are subsidiary to understanding what another is trying to say. Thus, to find a bridge, and to provide a more comprehensive account of meaning generation during interpersonal interactions, dynamical accounts must be extended to account for the effects of partner-specific knowledge in conversation.

To do so, we argue that attributional information serves the same function as perceptuomotor cues during communication. Instead of providing the means by which movement is coordinated, they act to constrain mutual understanding between language users. Thus, the attributes are a control parameter that influences perspective-taking during linguistic interpretations. This notion of control is analogously found in joint action tasks where perceptual affordances guide cooperative behavior. In Richardson, Marsh, and Baron (2007), people moving planks of wood have been shown to predictively switch from autonomous to cooperative action based on a relationship between each other's arm span and the length of the plank. In similar fashion, language users will take on a particular interpretative stance, that is more or less cooperative in establishing shared meaning, based on "affordances," or attributions, that are rapidly assimilated and reinforced throughout communicative interactions. These attributions act to warp comprehension processes from the start, influencing how individuals come to exhibit stable, yet flexible patterns of responding.

To relate this dynamical process to human response behavior in a linguistic task, we turn to a simple dynamical model of experimental data. This model is derived from a prominent mathematical model of bimanual motor coordination. Originally developed by Haken, Kelso, and Bunz (1985) (HKB) to capture phase transitions in what is called a "bistable attractor landscape" (explained further below), this model has been extended to a variety of domains, revealing widespread commonalities between perceptual, cognitive, and motor systems (e.g., Engstrom, Kelso, & Holroyd, 1996; Frank, Richardson, Lopresti-Goodman, & Turvey, 2009; Tuller, Case, Ding, & Kelso, 1994; van Rooij, Bongers, & Haselager, 2002; see Chemero, 2009; Schmidt & Turvey, 1995; for reviews). The value of this dynamical model is that it captures complex behaviors based on simple, unifying principles of behavioral change brought about by situated, environmental constraints.

In the HKB extension for our perspective-taking task, we follow a strategy that adheres to basic steps laid out in previous research (Beer, 2003; Raczaszek-Leonardi & Kelso, 2008). First, we need to find a tractable way of expressing the coupling between language users in a communicative task. This requires reducing the multiple sources of information involved in an interaction (i.e., system complexity) to a quantifiable and transparent outcome variable. By doing so, this simple behavioral variable can then be used to characterize cognitive processing in the interactive task. Second, we need to identify the parameters that constrain (or govern) how this cognitive process emerges or changes. Third, we must develop a version of the dynamic model under these constraints and show how its behavior maps onto human performance, thus providing a qualitative demonstration of the unfolding dynamics observed in that performance.

The human data we model is taken from a recently published study of Duran, Dale, and Kreuz (2011). In this task, participants were required to interpret verbal instructions from a seemingly real, but simulated partner who directed them to select an object on a computer screen. Occasionally, instructions could be ambiguous with respect to which object (e.g., one on the left, or the other on the right) should be selected. Although language users were not physically co-present, the spatial referent was ostensibly visible to both. Depending on attributional information available about their computer partner, participants either grounded interpretation from their own visual perspective (i.e., egocentric stance), or from the visual perspective of their partner (i.e., "other-centric" stance). In other words, an ambiguous description could be resolved as a selection indicating *perspective*: "choosing the object on *my* left" vs. "choosing the object on *their* left." In terms of our simulation strategy, these interpretative stances constitute behavioral outcomes captured with a one-dimensional variable. Obviously perspective-taking is based on a diverse range of requisite cognitive processes, but for current purposes, outcome is expressed on a single dimension: Which perspective is the participant taking? This low-dimensional characterization is in terms of egocentric vs. other-centric *response distributions* as indicating which stable