Cognition 180 (2018) 38–51

Contents lists available at ScienceDirect

Cognition

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

Communicative intent modulates production and comprehension of actions and gestures: A Kinect study

James P. Trujillo[a,b,*], Irina Simanova[a], Harold Bekkering[a], Asli Özyürek[b,c]

[a] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands
[b] Centre for Language Studies, Radboud University Nijmegen, The Netherlands
[c] Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525XD Nijmegen, The Netherlands

ARTICLE INFO

Keywords:
Pantomime
Communication
Multi-modal
Motion capture
Eye-gaze
Action

ABSTRACT

Actions may be used to directly act on the world around us, or as a means of communication. Effective communication requires the addressee to recognize the act as being communicative. Humans are sensitive to ostensive communicative cues, such as direct eye gaze (Csibra & Gergely, 2009). However, there may be additional cues present in the action or gesture itself. Here we investigate features that characterize the initiation of a communicative interaction in both production and comprehension.

We asked 40 participants to perform 31 pairs of object-directed actions and representational gestures in more- or less- communicative contexts. Data were collected using motion capture technology for kinematics and video recording for eye-gaze. With these data, we focused on two issues. First, if and how actions and gestures are systematically modulated when performed in a communicative context. Second, if observers exploit such kinematic information to classify an act as communicative.

Our study showed that during production the communicative context modulates space–time dimensions of kinematics and elicits an increase in addressee-directed eye-gaze. Naïve participants detected communicative intent in actions and gestures preferentially using eye-gaze information, only utilizing kinematic information when eye-gaze was unavailable.

Our study highlights the general communicative modulation of action and gesture kinematics during production but also shows that addressees only exploit this modulation to recognize communicative intention in the absence of eye-gaze. We discuss these findings in terms of distinctive but potentially overlapping functions of addressee directed eye-gaze and kinematic modulations within the wider context of human communication and learning.

1. Introduction

Our hands may be used in a variety of ways to interact with the world around us. Two such interactions are object-directed actions, in which the hands interact with a physical object (e.g., to open a jar), and representational gestures (Kendon, 2004; McNeill, 1994), in which the hands are used to simulate an interaction or visually represent a non-present object (hands move as if opening a jar). What is specific to humans is that both categories of movements can be recruited for the purpose of communication, allowing us to teach through demonstration (Campisi & Özyürek, 2013; Southgate, Chevallier, & Csibra, 2009) or convey the intention for an observer to act in response (Tomasello, 2010).

Characteristic of communicative acts is the accompanying addressee-directed eye-gaze (Brand, Shallcross, Sabatos, & Massie, 2007). Humans in particular seem inherently sensitive to ostensive communicative cues, such as direct eye gaze and eyebrow raise (Csibra & Gergely, 2009). Direct eye-gaze is particularly powerful, displaying a willingness to interact (Cary, 1978), as well as altering cognitive processing and behavioural response (Senju & Johnson, 2009). For example, a recent study by Innocenti et al. investigated the impact of eye-gaze on a requesting gesture, e.g. reaching out and grasping an empty glass with the implied request to have it filled. The study showed that both the speed and size of a communicative gesture and addressee-directed eye-gaze affected kinematics of the response act. Therefore, the mere presence of direct eye-gaze induced a measurable effect on the response of the addressee (Innocenti, de Stefani, Bernardi, Campione, & Gentilucci, 2012).

For communication in general, there are at least two main requirements: the communicator must make his or her intention to

* Corresponding author at: Radboud University, Montessorilaan 3, B.01.25, 6525GR Nijmegen, The Netherlands.
E-mail address: jptrujillo88@hotmail.com (J.P. Trujillo).

https://doi.org/10.1016/j.cognition.2018.04.003
Received 24 January 2017; Received in revised form 16 March 2018; Accepted 2 April 2018

communicate recognizable, and they must represent the semantic information they wish to be received by the observer. The first step in communicating using actions or gestures is thus for the communicator to make the action or gesture recognizable as being a communicative act. In doing so the communicator might use kinematic modulation (see, for example, (Becchio, Cavallo, et al., 2012)) as well as addressee-directed eye-gaze (Kampe, Frith, & Frith, 2003; Schilbach et al., 2006). Secondly the communicator's cues need to be picked up by addressee in order to interpret actions or gestures as communicative. Here, again, both the kinematics of the manual acts and the ostensive cues, or the interaction of both, can play a role. In the present study, we address the overall profile of communicative actions and gestures within the larger context of production and comprehension. We compare for the first time actions and gestures in more-communicative versus less communicative contexts to see if they are subject to similar kinematic modulations and are coupled by ostensive cues. We then investigate whether and how these cues are in turn interpreted by addressees. To quantify kinematic modulation effects we use the Kinect device to obtain a non-intrusive, objective and precise measure of action and gesture. The next few paragraphs summarize the current literature on the kinematic modulation and on the perception of actions and gestures in communicative context.

2. Production of communicative actions and gestures

At the basic motor control level, actions are thought to follow a principle of motor efficiency (Todorov & Jordan, 2002). In this framework, control of an action is a balance between reducing cost and achieving the goal of the action. While this framework explains action control in a neutral setting, there is evidence that other contextual or cognitive domains influence these dynamics. The intention to communicate affects the velocity of reach-to-grasp movements (Sartori, Becchio, Bara, & Castiello, 2009), and can modulate the trajectory of such movements to make a target more predictable to a co-actor (Sacheli, Tidoni, Pavone, Aglioti, & Candidi, 2013). Furthermore, child-directed communicative actions are marked by several kinematic modulations, including an increased range-of-motion and punctuality (Brand, Baldwin, & Ashburn, 2002). At the level of cognitive and neural implementation of motor control, this indicates a top-down influence on action production that is theorized to facilitate interactions by balancing the initial efficiency principle with the additional factor of disambiguating the end-goal for an observer (Pezzulo, Donnarumma, & Dindo, 2013). In line with the account by Pezzulo and colleagues, we suggest that the kinematic modulation from a communicative context can be summarized as an optimization of space-time dimensions (Pezzulo et al., 2013). In this account, communicative modulation is an effort to present the optimal amount of visual information to disambiguate the act (optimization of space) within an efficient amount of time (optimization of time). We extend this framework by investigating specific kinematic cues, and testing how ostensive eye-gaze is implemented together with kinematic modulation in both actions and gestures. As actions are not inherently communicative, and indeed less likely to be interpreted as communicative by observers (Kelly, Healey, Özyürek, & Holler, 2015; Novack, Wakefield, & Goldin-Meadow, 2016), it may be that direct eye-gaze is an important communicative cue for actions. An additional open question is whether similar communication modulations occur not only in actions, but also in representational gestures.

Although the motor efficiency/optimization principle does not specifically refer to gestures, they too are manual acts with a specific extrinsic goal. Often, this goal to change the internal state of an observer, but gestures may also be performed without communicative intention. For instance, in the context of co-thought gestures, one uses gestures while trying to solve complex visuospatial tasks (Chu & Kita, 2011). Additionally, clinicians often use pantomime production tasks as a clinical measure in aphasia (Goldenberg, Hartmann, & Schlott, 2003; Hermsdörfer, Li, Randerath, Goldenberg, & Johannsen, 2012). Gestures then are likely to also follow an initial efficiency principle which may further be modulated depending on the goal or intention. Like actions, gestures are also influenced by a communicative context. For example, when meant to be more informative to an observer, pointing gestures are made slower than when the gesture will not be used by an observer (Peeters, Chu, Holler, Hagoort, & Özyürek, 2015). Furthermore, during a demonstration or explanation, a gap in common knowledge between speaker and addressee leads to gestures that are larger (Bavelas, Gerwing, Sutton, & Prevost, 2008; Campisi & Özyürek, 2013), more complex or precise (Galati & Brennan, 2014; Gerwing & Bavelas, 2004; Holler & Beattie, 2005) and are produced higher in space (Hilliard & Cook, 2016). Whether these kinematic modulations are comparable to those observed in actions in similar communicative settings, has not been assessed. This is of interest because gestures are reliant on kinematics to convey meaning, whereas actions can utilize the object (manipulation) to convey meaning. We could then expect the two modalities to differ in the way they are made more communicative. For example, because gestures are more inherently communicative, the strong direct eye-gaze signal may be less important for gestures compared to actions. Therefore, an interesting open question is whether the same kinematic and eye-gaze features are modulated when the two modalities are performed in a more communicative context.

3. Perception of communicative actions and gestures

Although communicative intent driven modulation is present during the production of actions and gestures, as shown above, it is less clear whether and how this modulation is seen or used by observers. Studies show that children prefer actions marked by increased range of motion and exaggerated movement boundaries (Brand et al., 2002), which leads to increased visual attention in infants (Brand & Shallcross, 2008), and more frequent imitation of a demonstrated action in children (Williamson & Brand, 2014). In regard to intention recognition, a study on social actions by Manera et al., showed that observers are able to distinguish between cooperative and competitive actions using only the kinematics (point-light-displays) (Manera, Becchio, Cavallo, Sartori, & Castiello, 2011). This suggests that kinematic modulation, at least in regard to child-directed actions and social context, is noticed by observers.

With regard to perception of the communicativeness of gestures, a recent study by Novack et al. shows that movements in the presence of objects are seen as representations of actions, while the *same* movements made in the absence of objects are described as being movement for its own sake (Novack et al., 2016). This suggests that even though kinematics clearly affects the way the action or gesture is perceived, observers rely strongly on situational constraints to understand the underlying intention. Further evidence comes from a study on body orientation and iconic gesture use (Nagels, Kircher, Steines, & Straube, 2015). Nagels and colleagues found that when a speaker is oriented toward an addressee and gestures during speech, the addressee feels more addressed, thereby indicating a better recognition of communicative intent. Interestingly, both the condition with the speaker orientated towards the addressee but not using iconic gestures as well as the condition with the speaker oriented away from the addressee but using iconic gesture were also rated as being more communicative than the condition in which the speaker faced away and did not use gestures (Nagels et al., 2015). These studies indicate that, at least for iconic gestures, both eye-gaze directed to the addressee and gestures can convey a communicative intent. It is important to note that although iconic gesture use contributed to the feeling of being addressed, the kinematics of gestures themselves were not modified in that study. To date, there are no studies that have investigated kinematic modulation of gestures in comprehension. Therefore the question remains of how such a modulation will impact the perceived communicativeness of the gesture or the action. Furthermore, previous research comparing the