Title
When Coordination is Worth a Thousand Words: the Role of Gesture in Grounding

Permalink
https://escholarship.org/uc/item/9k5080ch

Journal

ISSN
1069-7977

Authors
Appelbaum, Meredyth Krych
Schultheiss, Joan
Banzon, Julie

Publication Date
2004

Peer reviewed
When Coordination is Worth a Thousand Words:  
the Role of Gesture in Grounding

Meredyth Krych Appelbaum (krychm@mail.montclair.edu)  
Joan Schultheiss (schultheisj1@mail.montclair.edu)  
Julie Banzon (banzonj1@mail.montclair.edu)  
Department of Psychology, Montclair State University  
1 Normal Avenue, Montclair, NJ 07043 USA

Introduction

In most conversations, people rely on a process of grounding, in which people establish the mutual belief that they have been understood (Clark, 1996). While the majority of research on grounding has focused on speech in conversation, we examine the grounding process when people may coordinate only through gesture, compared to speech + gesture, and to speech-only conditions.

Grounding is fundamentally about coordination between people. Clark & Krych (2004) demonstrated how speaking and listening are incremental processes and how many of those increments are determined jointly—whether through speech or gesture.

In this study, our aim is to achieve a better understanding of the role of gestures and speech in communication and what gesturing alone can further inform us of the grounding process.

Lozano and Tversky (2003) have studied gestures and speech+gestures when people are videotaped assembling a TV cart for an undefined audience who would later view the videotape. In our research, we are interested in how two participants interact and coordinate with one another under different conditions.

While one might predict that not being allowed to talk with one another would be a disadvantage, resulting in more time to achieve understanding, we predict that the gesture-only and speech + gesture conditions should be equivalent in timing, while the speech-only condition should take significantly longer.

Methods

Pairs of students worked together as one participant, the director, instructed the other participant, the builder, how to create duplicate models of Lego blocks. Their goal was for the builder to create identical models as efficiently as possible based on the director’s instructions. Each pair had a practice trial to orient them to the task and then constructed nine other models. The models used were the same as in Clark & Krych (2004) and Krych & Clark (1997).

Thirty-nine subject pairs participated in one of three separate conditions. Each condition consisted of 13 subject pairs who were all undergraduate students. In one condition, the director could see the builder’s workspace and they could converse normally using both speech and gesture as they wished (speech + gesture). In a second condition, the builder’s workspace was not visible to the director, so participants could only communicate with speech (speech-only). A third group of subject pairs participated in a gesture-only condition in which the workspace was visible to the director, but the subjects could not use any words at all. They could communicate only by gesturing to one another and pointing to objects.

Results and Discussion

As predicted, there was a large difference in the average amount of time to complete each model, F(2, 36) = 27.64, p< .001). The speech-only condition took much longer— 181 seconds compared to 94.5 seconds in the Speech + Gesture condition and 112 seconds in the Gesture-only condition. The latter two conditions were statistically equivalent to one another. This pattern held true even if the practice trial was included. Thus, participants who were restricted to gesturing were not at a disadvantage compared to participants who could speak and gesture. These results appear consistent with the findings of Lozano and Tversky’s non-interactive study (2003) that language and gesture can supplement as well as complement each other.

In the future, we plan to focus on the process of how people ground information in the gesture-only condition. We suggest that studying the process of grounding when people may only gesture to one another will shed further information on the underlying processes involved in achieving understanding in face-to-face conversation.

Acknowledgments

We would like to acknowledge Montclair State University for funding some of this research. We also thank Joanna Musial and Cecilia Sullivan for their dedicated work.

References