UC Merced

Title

Gesture Production and Theory of Mind:Effective Disambiguation in Communication through Gesture

Permalink

Journal

Authors

Kim, Minju
Schachner, Adena

Publication Date

2020

Peer reviewed

Gesture Production and Theory of Mind:
Effective Disambiguation in Communication through Gesture

Minju Kim (minjukim@ucsd.edu), Adena Schachner (schachner@ucsd.edu)
University of California, San Diego, Department of Psychology
9500 Gilman Drive, La Jolla, CA, 92092 USA

Abstract

People design their speech acts with their listeners in mind, accounting for their knowledge and other mental states. Is this ability specific to spoken language and co-speech gesture, or does it appear in pantomimic gestures as well? We ask whether adults flexibly shift their silent gestures to emphasize relevant information, representing different features of the target in different contexts. In a two-item reference game, adults gestured to a partner to indicate which object was the target. Item pairs differed in one of three features (size, shape, pattern). We found that adults were more likely to gesture a feature when it was relevant to distinguishing the two possible referents, versus when it was not. Thus, adults flexibly modified their gestures to meet their partners' needs, emphasizing the relevant feature. These data lay a foundation for future work on the development of use of theory of mind in gestural communication in childhood.

Keywords: gesture production; referential communication; theory of mind; common ground; disambiguation; language

Introduction

People face the constant challenge of communicating effectively, across many different contexts and interlocutors (Nadig & Sedivy, 2002; Wardlow, 2013). For communication in speech, theory of mind (reasoning about others' mental states) plays a major and important role (Wellman, 2013; Wellman & Liu, 2004). For example, common ground between two interlocutors dramatically affects the utterances people produce, showing that adults use social reasoning to shape their linguistic communication (Clark, 1996; Clark & Marshall, 1981). More generally, effective communication requires reasoning about the mental states of the other person, to determine what the other person will understand, and what pragmatic inferences they might make (Frank & Goodman, 2012; Watson et al., 1999).

Theory of mind in speech has been formalized as production and interpretation of rational speech acts (Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013). According to rational speech-act theory, listeners assume that speakers will produce utterances using two assumptions of Gricean reasoning (Grice, 1975): that the speakers are approximately rational, and that the speakers are cooperative in optimizing their utterances for the communicative goal. Speakers use these Gricean maxims to determine not only what to say, but also what not to say. For example, when writing a recommendation letter, if the writer does not have positive things to say about a typical topic, they may choose not to mention that topic; and expect that the reader will 'read between the lines.'

When using speech to distinguish between multiple possible referents, speakers use Gricean logic to select words that not only describe the target, but allow the speaker to identify it from the set of other items. Thus, if a target referent has two features, one of which applies to multiple possible referents, and one of which is unique, people will choose to state the unique feature (Frank & Goodman, 2012). In the current paper, we explore whether this kind of pragmatic reasoning shapes production when communicating using gesture, rather than speech.

Co-speech Gestures and Theory of Mind

Verbal communication is often accompanied by co-speech gestures (Kendon, 2004; McNeil, 1992). Together, speech and co-speech gestures function as an integrated system of communication (Kendon, 2004), in which gestures are co-expressive and synchronous with speech (Goldin-Meadow et al., 1996).

There is evidence that theory of mind shapes the production of co-speech gestures, maximizing the effectiveness of communication. When used along with speech, gestures result in more effective communication than speech alone (Clark, 2016; Iverson & Goldin-Meadow, 1997). When producing co-speech gestures, people spontaneously use gestures to resolve ambiguity in speech, disambiguating the linguistic referent from other potential referents (e.g. iconic gestures in narratives; Dick et al., 2014). Speakers also use more co-speech gestures when the listener needs more information, and fewer co-speech gestures when the listener already has access to the information (Galati & Brennan, 2014; Jacobs & Garnham, 2007). Similarly, speakers use smaller and less precise gestures when the listener has versus has not heard a story before (Galati & Brennan, 2014), and gesture more when the listener has visual access to them (Hoetjes et al., 2015). Thus, when people produce spontaneous co-speech gestures alongside natural speech, gesticulation differs by the semantic need for disambiguation and the knowledge state and perceptual access of the listener, requiring theory of mind.

Silent Gestures and Symbolic Communication

Co-speech gestures contrast with silent (un-voiced, or pantomimic) gestures. Unlike co-speech gestures, silent gestures function alone, and thus "carry the full burden of communication" (Goldin-Meadow et al., 1996). When

hearing people are restricted from using speech and are instructed to only use silent gestures, they can successfully describe events, and their gestures furthermore show some grammatical features of conventional language (e.g. Bohn, Kachel, & Tomasello, 2019; Fay et al., 2014; Goldin-Meadow et al., 1996; 2008). For example, people asked to describe a series of events either only using gestures (vs. using speech, with no instruction regarding gestures) segmented and hierarchically combined their gestures, as in the grammar of conventional language. Silent gestures also showed more distinctive features and were less restricted in their forms than co-speech gestures (Goldin-Meadow et al., 1996). These gestures show some features similar to the home-signs created by deaf children of hearing parents, where the children are not regularly exposed to a conventional sign language and use the manual modality to communicate (Goldin-Meadow & Mylander, 1990).

The Current Study

Do people use theory of mind to communicate effectively and rationally when using silent gestures? That is, when selecting between multiple possible referents, do they selectively use gestures that not only describe the target, but distinguish the target from other possible referents? Our goal in the current study is to characterize adults' ability to flexibly change their silent gesture to emphasize the relevant information in different referential contexts. Using a simplified version of minimal language game for referential communication (e.g. Frank & Goodman, 2012), we probe how people flexibly shift their gesticulation to represent different features of the target, so as to emphasize the relevant information in each particular referential context.

While there have been relevant previous studies regarding co-speech gesture, studies have not explored the use of gesture separate from speech; and also used a different manipulation, changing what information was needed by manipulating the identity of the partner, and thus their knowledge states (Galati & Brennan, 2014; Jacobs & Garnham, 2007). In our study we had each participant stay with one partner, and created the need to flexibly shift what should be communicated by changing the pairings of stimuli, modifying which of three different features distinguished the target from the distractor (size, shape, or pattern).

We aimed to create a simple task that can be used with young children in the future, to explore the development of the use of theory of mind in gestural communication. We predict that adults will be able to use theory of mind to produce effective gestural communication, modifying their gestures to communicate the relevant feature for distinguishing the target from the distractor. Regarding children's performance, two accounts make opposing predictions. Recent work has suggested that reference may be easier to establish in gesture than vocalization (Fay et al., 2014). This leads to the prediction that even children would succeed, since children can modify their speech to take into account the listener's perspective (Nilsen & Graham, 2009; Wardlow & Heyman, 2016). However, there are still limitations to children's abilities even at school-age (Girbau, 2001; Sonnenschein & Whitehurst, 1984; Whitehurst & Sonnenschein, 1982), and at preschool age, children show a reduced tendency to spontaneously use gesture in a communicative way, requiring prompting from an adult (Bohn, Kachel, & Tomasello, 2019), leading to the prediction that children may fail. Thus, the current paper aims to provide a measure of adult performance, to allow for future comparison and exploration of the development of use of theory of mind in gestural communication.

Using a two-item reference game with pairs of participants, people were asked to use gestures (not speech, and excluding pointing) to indicate to a partner which of two items was the target referent, similar to a game of charades. The responder watched the gesturer on a video screen, and was asked to choose which of the two items was the intended target, based on what they saw.

Crucially, we manipulated how the target object differed from the non-target object: For each pair, the target differed on one feature, but was identical on the other two features (size, shape, or pattern). We asked whether people were more likely to produce gestures corresponding to each feature (e.g. size) when it was relevant to distinguishing the two possible referents (e.g. a big-striped-pointy object and a small-striped-pointy object), versus when it was not relevant for that context (e.g. a big-striped-pointy object and a big-dotted-pointy object). If so, this would provide evidence that adults are able to tailor their silent gestures to provide the particular information the observer needs, given the current communicative context.

Methods

Participants

N=50 people participated, in 25 pairs (M_{age} = 20.33 years, SD_{age}=2.45; 13 males). Participants were recruited from the student population at a large public university in Southern California, and earned course credit in exchange for participation. The study was approved by the university IRB. Two additional pairs of participants participated but were excluded due to producing vocalizations in addition to gestures (1 pair), or due to experimenter error (1 pair).

Stimuli & Design

Stimuli consisted of eight novel geometric objects, designed to differ on three features: Shape (pointy vs. round), size (big vs. small) and pattern (stripes vs. dots; see Figure 1a).

On each trial, a pair of two shapes were presented next to one another on a computer screen. Within each pair, the two items differed on one feature (e.g., size) and were the same on the two other features (e.g. shape, pattern). Thus, for each target item, three test pairs were created, pairing the target item with a contrast item that differed in either size, shape, or pattern. This made for a total of 24 unique test trial pairs, all tested within-subject. The order and location of the target

item (left, right) was counterbalanced across subjects, and the order of trials was pseudorandomized.

Procedure

Participants were tested in the lab, seated on chairs in the same room on opposite sides of an opaque black curtain.

One participant served as the gesturer. On the gesturer's side of the curtain was a computer which displayed the stimuli, with the addition of a red arrow indicating which of the two objects was the target (see Figure 1b). The gesturer was instructed to describe the object under the red arrow with gestures, to indicate to their partner which object was the target. They were asked to do this as quickly as possible and using as few gestures as possible. The gesturer was further instructed not to use speech, not to point to the screen, and to gesture only about the target item. Gestures were captured using a video camera (Panasonic HC-V180K), set up to capture a wide area surrounding the gesturers' upper body, including the space above and around their head and torso, down to their lap or the edge of the chair they were sitting on. The other participant served as the responder. On the responder's side of the curtain was a large TV, on which they viewed the gesturer on a live video feed at approximately life-size (49-inch LG 49LH570A). In a web browser on a computer (iMac, using an online survey created with Qualtrics software), the responder viewed the pairs of stimuli and could click to record their answers regarding which was the target item.

Participants first completed three practice trials, using different stimuli from the main test trials (a 5-point star vs. a heart shape; a big circle vs. a small circle; a square with horizontal zigzag patterns vs. a square with spiral patterns). Before starting the main trials, participants were shown an array of all eight possible items, and were told that pairs of items could differ in shape, size, or pattern (see Figure 1a).

On each trial, both participants saw the same pair of objects on separate computer screens; only the gesturer saw a red arrow indicating which was the target item. The gesturer then used gestures to indicate to their partner which was the target item, with a maximum time available of 8 seconds. When the gesturer was done producing gestures, they pushed a button to indicate this to the responder. Then, the responder silently judged which item was the target (by clicking one of the two answers on their computer; the gesturer did not receive feedback or learn which one the responder had chosen). After submitting their answer, the responder indicated they were done by ringing a bell; at which point the experiment moved on to the next trial. The entire session took about 18 minutes to complete.

Gesture Coding

All gestures were coded from video by a coder who was blind to the counterbalance conditions regarding the display (e.g. target on the left or right), or the sequence of the object pairs presented. The coder was unaware of which stimuli were being presented to the gesturer and the responder, and were
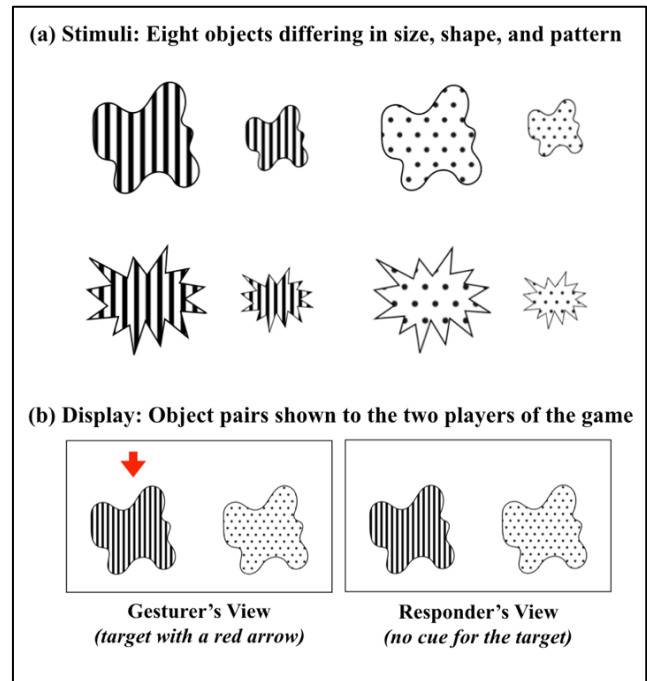

Figure 1. Stimuli and method. On each trial, the gesturer and responder saw a pair of objects, and the gesturer used movements (not speech, not pointing) to communicate to which item was the target. (a) Stimulus pairs each differed on only one feature (size, shape, or pattern). (b) On two screens, both people saw the same two objects; a red arrow indicating the target object was shown only to the gesturer and not to the responder.

thus blind to the nature of both the target item and the contrast item.

Coders were asked to categorize each gesture by what feature or features were being represented (size, shape, or pattern). A single gesture could be coded as representing multiple features (e.g. size and shape). Gestures that were ambiguous were coded as such; if coders had a guess of what feature was being represented by an ambiguous gesture, this coding was included in analysis. For each trial, coders also recorded the number of gestures and the order in which multiple gestures were produced. To check the reliability of coding, we had two independent coders code a subset of the data (40% of participants; 10/25 pairs of participants), and compared the similarity of the two coders' categorizations across all trials. We found that coders agreed with one another 95.56% of the time (688 out of 720 features) regarding whether each feature was gestured.

Results

If participants flexibly change their gestures to highlight the relevant information on each trial, then the relevance of the feature should predict whether that feature is gestured, or not. To test this prediction, we constructed a logistic regression

model predicting whether or not the feature was gestured, with the following predictors: the type of feature (size, pattern, shape), whether or not the feature was relevant on that trial (relevance), the interaction between the feature type and relevance, and subject as a random factor. An initial test showed that the interaction term was not a significant predictor (nested model comparison, $\chi^2(2) = 2.65$, $p = 0.26$), and it was thus removed from the model in all subsequent analyses.

To ask whether participants were more likely to gesture a feature when it was relevant to distinguishing the two items versus when it was irrelevant, we performed a nested model comparison of the model with vs. without relevance as a predictor. We found that the features' relevance added significant predictive value to the model, such that the full model fit the data better than the simpler model with relevance removed (nested model comparison; $\chi^2(1)$ =873.77, $p < 0.0001$).

Thus, people's tendency to gesture a particular feature was affected by whether it was relevant to distinguishing the two items, or not. In particular, participants were more likely to gesture the feature when it was relevant than irrelevant (Z=19.87, $\beta = 4.79$, p<0.0001). This pattern held across each of the three features. On trials where size was relevant, people gestured size 87.0% of the time on average (SEM=5.02%), vs. 22.75% on trials where size was irrelevant (SEM=35.53%). Similar patterns were seen for pattern (Relevant: M=98.0%, SEM=2.0%; Irrelevant: M=37.75%, $SEM = 6.47\%$) and shape (Relevant: M=96.5%, SEM=2.45%; Irrelevant: M=41.25%, $SEM = 6.36\%$).

Feature type was also a significant predictor, such that people were less likely to gesture certain feature types than others (nested model comparison, full model vs. feature type removed; $\chi^2(1)$ =81.44, $p < 0.0001$). In particular, size was gestured less often than shape or pattern (Z= -7.11, β= -1.37, $p < 0.0001$).

Was the Gestural Communication Successful?

We also analyzed the responders' accuracy, to determine whether the gestural communication was successful or not. Overall, responders were highly successful in identifying the target item: They were correct on 98.83% of trials (23.72/24), with 20 out of 25 participants getting all trials correct, and a maximum of 2 incorrect trials for any responding participant.

We next asked whether the responder was less accurate on trials where the relevant gesture had not been produced, as would be expected since this information was the only way to disambiguate the two referents. However, the relevant gesture was only omitted on 37 out of 600 trials (6.1%), and for many participants, the relevant gesture was never omitted, making statistical comparison difficult. However, as predicted, the 37 trials without relevant gestures had lower responder accuracy than the 563 trials with relevant gestures (86.49% vs. 99.64%).

How was the responder able to answer correctly on 86% of trials, even without seeing the relevant gesture? This may be
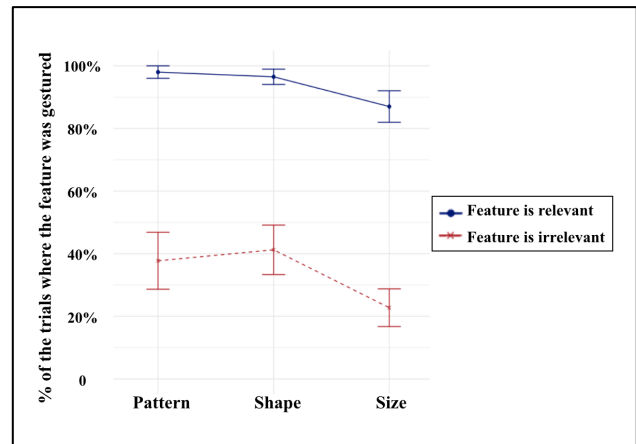

Figure 2. Proportion of trials on which gestures were produced by feature types. When a feature was relevant to the goal of distinguishing the two referents, gesturers were much more likely to produce that feature (in blue) than when it was irrelevant (in red). Error bars represent the standard error of the mean.

explained by the presence of gestures coded as 'ambiguous' in many of these trials (40.62%, 13 out of 32 trials). It is possible that these ambiguous gestures contained some relevant information; and that the responders were simply more sensitive to this information than the coders.

Discussion

Here we find evidence that when producing gesture, people tailor their silent gestures to make them maximally effective at communication given the context. In particular, depending on which feature was relevant to distinguishing two items, people changed how they represented that item in their gestures, focusing selectively on the feature that was relevant to distinguishing the two items. Thus, people appear able to reason about what information is needed to identify the target item, and adjust their gestural communication accordingly. This suggests that theory of mind can play an important role in shaping gestural communication, and speaks to parallels between communication in silent gesture and communication in language.

To achieve effective communication, speakers must balance the goals of efficiency and clarity, which may collide with one another (Fay & Ellison, 2013). Speakers maximize efficiency by being selective with the information that they provide, in order to use as few words or syllables as possible (Gibson et al., 2019). To do this, one must select the relevant information to minimize ambiguity within the current context (Wilkes-Gibbs & Clark, 1992). In our tasks, we find these same goals at work: People in our task maximized efficiency and clarity when producing gestures, by selectively representing relevant information and eliminating irrelevant information. Beyond speech and silent gesture, recent work has shown that adults can tailor their drawings to signal relevant information for the context (Mukherjee et al., 2019),

potentially suggesting that the use of theory of mind in referential communication is not necessarily constrained to our most commonly used modalities.

Parallels Between Gesture and Language

Overall, our findings support the idea that communication through gesture has deep parallels with conventional language (Goldin-Meadow et al., 1996; Goldin-Meadow et al., 2008). These findings are in line with the idea that gesture can quickly transform to reflect the systematic features of spoken language: Children intuitively and rapidly transform silent gestures into a conventional system, when the need arises (e.g. Nicaraguan sign language; Kegl et al., 1999; Senghas & Coppola, 2001; see also Bohn, Kachel, & Tomasello, 2019). Recent work has suggested that reference may be easier to establish in gesture than vocalization (Fay et al., 2014), prompting others to theorize that human conventional communication may have originated in gesture (Bohn, Call, & Tomasello, 2019).

Our findings fit into this general theoretical picture, in which gestural communication (even by non-signers) bears deep similarity to language, rather than fundamental differences originating from conventionality of communication modality. In future studies, we hope to explore whether and how people flexibly shift from language to gesture, to intuitively use whichever modality makes it easiest to represent the information they are trying to communicate (e.g. it may be impossible to gesture color, while an unfamiliar shape may be easier to communicate through gesture than spoken word).

The Development of Gestural Communication

These data also lay a foundation for developmental work on children's use of theory of mind in gestural communication. A substantial literature has explored the impact of gesture on children's learning: When children are instructed to use gestures, they perform better on math problems (Goldin-Meadow et al., 2009; Novack et al., 2014) and mental rotation tasks (Wakefield et al., 2019). Children's spontaneous gestures can also show implicit understanding of a concept, even when explicit verbal answers are incorrect, signalling a transitional state of understanding predictive of future learning (Church & Goldin-Meadow, 1986; Goldin-Meadow & Wagner, 2005; Perry et al., 1988). However, few studies have examined children's ability to use gesture to communicate in a social context. Observing children's gesture production in referential communication may shed light on children's problem-solving abilities in communication, which may have been underestimated due to limitations in verbal ability with cognitive overload (e.g. structuring one's thoughts into full sentences; Simon & Hayes, 1986).

For young children, theory of mind develops extensively over the preschool and early school years, raising the question of whether children's ability to tailor their utterances for effective communication changes over this period as well. Theory of mind clearly enters into communication from an early age in some ways: As early as two, children succeed in guiding their caregivers' attention to a desired object (O'Neill, 1996), and 5- to 6-year-olds infer two others' mutual knowledge in a production-elicited task (e.g. Nadig & Sedivy, 2002).

However, even school-age children show dramatic differences from adults: Six-year-old children show difficulty in gauging the amount of information needed, and in selecting the disambiguating information their partner needs (Sonnenschein & Whitehurst, 1984; Whitehurst & Sonnenschein, 1982). Even at age 10, although children are well aware of the disambiguating information between multiple referents, they still often provide ambiguous information in a communicative interaction (Girbau, 2001).

Recent data also suggests that at preschool age, children show a reduced tendency to spontaneously use gesture in a communicative way, requiring prompting from an adult (Bohn, Kachel, & Tomasello, 2019). In future work, we plan to compare children's ability to tailor their gestures for their listeners' needs with their abilities to do this via spoken language. For children, unlike adults, there may be a dissociation such that children more flexibly change their spoken utterances in response to the communicative requirements than their gestures -- perhaps due to a greater amount of experience and practice with conventional communication. Alternatively, recent findings and theories regarding the primacy of gesture over spoken language (Bohn, Call, & Tomasello, 2019; Fay et al., 2014) motivate the contrasting prediction that children may actually be more successful in tailoring their gestural communication for their listeners' needs than their spoken conventional communication.

Conclusion

Overall, we find that when producing silent gestures, people flexibly shift which features they represent or omit from their gestures to emphasize the relevant, disambiguating information needed by the listener in each context. This suggests that adults design their silent gestures for the listeners' needs, and that theory of mind plays a similar role in communication in silent gesture as in spoken language. These data lay a foundation for future developmental work on children's use of theory of mind in gestural communication.

References

Bohn, M., Call, J., & Tomasello, M. (2019). Natural reference: A phylo- and ontogenetic perspective on the

comprehension of iconic gestures and vocalizations. *Developmental Science, 22*, e12757.

Bohn, M., Kachel, G., & Tomasello, M. (2019). Young children spontaneously recreate core properties of language in a new modality. *Proceedings of the National Academy of Sciences, 116*, 26072-26077.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*, 43-71.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H. (2016). Depicting as a method of communication. *Psychological Review, 123*, 324-347.

Clark, H. H., & Marshall, C. R. (1981). Definite Knowledge and Mutual Knowledge. In A. K. Joshi, B. L. Webber & I. A. Sag (Eds.), *Elements of Discourse Understanding*. Cambridge, UK: Cambridge University Press.

Dick, A. S., Mok, E. H., Beharelle, A. R., Goldin-Meadow, S., & Small, S. L. (2014). Frontal and temporal contributions to understanding the iconic co-speech gestures that accompany speech. *Human Brain Mapping, 35*, 900-917.

Fay, N., & Ellison, T. M. (2013). The cultural evolution of human communication systems in different sized populations: usability trumps learnability. *PloS one, 8*, e71781. doi:10.1371/journal.pone.0071781.

Fay, N., Lister, C. J., Ellison, T. M., & Goldin-Meadow, S. (2014). Creating a communication system from scratch: gesture beats vocalization hands down. *Frontiers in Psychology, 5,* 354.

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*, 998-998.

Galati, A., & Brennan, S. (2014). Speakers adapt gestures to addressees' knowledge: Implications for models of co-speech gesture. *Language, Cognition and Neuroscience, 29*, 435–451.

Gibson, E., Futrell, R., Piandadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences, 23*, 389-407.

Girbau, D. (2001). Children's referential communication failure: The ambiguity and abbreviation of message. *Journal of Language and Social Psychology, 20*, 81–89.

Grice, H. (1975). Logic and grammar. In D. Davidson & G. H. Harman (Eds.), *Readings in language and mind.* (pp. 64–75). Encino, CA: Dickenson.

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science, 20*, 267-272.

Goldin-Meadow, S., McNeill, D., & Singleton, J. (1996). Silence is liberating: removing the handcuffs on grammatical expression in the manual modality. *Psychological review, 103*, 34-55.

Goldin-Meadow, S., & Mylander, C. (1990). Beyond the input given: The child's role in the acquisition of language. *Language, 66*, 323-355.

Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences, 105*, 9163-9168.

Goldin-Meadow, S., & Wagner, S. M. (2005). How our hands help us learn. *Trends in Cognitive Sciences, 9*, 234-241.

Goodman, N. D., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science, 5*, 173-184.

Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E., & Swerts, M. (2015). Reduction in gesture during the production of repeated references. *Journal of Memory and Language*, 79, 1-17.

Iverson, J. M., & Goldin-Meadow, S. (1997). What's communication got to do with it? Gesture in children blind from birth. *Developmental Psychology, 33*, 453-467.

Jacobs, N., & Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language, 56,* 291–303.

Kegl, J., Senghas, A., & Coppola., M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. Degraff (Ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development.* Cambridge, MA: MIT Press.

Kendon, A. (2004). *Gesture: Visible action as utterance.* Cambridge, UK: Cambridge University Press.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

Mukherjee, K., Hawkins, R., & Fan, J. (2019). Conveying semantic part information in drawings. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society.*

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science, 13*, 329-336.

Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: Using the hands to learn math. P*sychological Science, 25,* 903-910.

Nilsen, E., & Graham, S. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology, 58*, 220–249.

O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledges state when making requests. *Child Development, 67*, 659-77.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359-400.

Senghas, A., & Coppola, M. (2001). Children Creating Language: How Nicaraguan Sign Language Acquired a Spatial Grammar. *Psychological Science, 12*, 323–328.

Simon, H. A., & Hayes, J. R. (1976). The understanding process: Problem isomorphs. *Cognitive Psychology, 8*, 165-190.

Sonnenschein, S., & Whitehurst, G. J. (1982). The effects of redundant communications on the behavior of listeners:

Does a picture need a thousand words? Journal of Psycholinguistic Research, 11(2), 115–125.

Sonnenschein, S., & Whitehurst, G. J. (1984). Developing referential communication: A hierarchy of skills. *Child Development, 55*, 1936–1945.

Wakefield, E. M., Foley, A. E., Ping, R., Villarreal, J. N., Goldin-Meadow, S., & Levine, S. C. (2019). Breaking down gesture and action in mental rotation: Understanding the components of movement that promote learning. *Developmental Psychology*, 55. 981-993.

Wardlow, L. (2013). Individual differences in speakers' perspective taking: The roles of executive control and working memory. *Psychonomic Bulletin & Review, 20*, 766-772.

Wardlow, L., & Heyman, G. D. (2016). The roles of feedback and working memory in children's reference production. *Journal of experimental child psychology, 150*, 180-193.

Watson, A. C., Nixon, C. L., Wilson, A., & Capage, L. (1999). Social interaction skills and theory of mind in young children. *Developmental Psychology*, *35*, 386.

Wellman, H. M. (2013). Universal social cognition. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us*. New York, NY: Oxford University Press.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*, 523-541.

Whitehurst, G. J., & Sonnenschein, S. (1982). The development of informative messages in referential communication: Knowing when versus knowing how. In W. P. Dickson (Ed.), *Children's oral communication skills.* New York, NY: Academic Press.

Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language*, *31*, 183-194.