John Benjamins Publishing Company

This is a contribution from *Integrating Gestures. The interdisciplinary nature of gesture*.
Edited by Gale Stam and Mika Ishino.
© 2011. John Benjamins Publishing Company

This electronic file may not be altered in any way.
The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.
Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.
For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).
Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

CHAPTER 4

Dyadic evidence for grounding with abstract deictic gestures

Janet Bavelas, Jennifer Gerwing, Meredith Allison, and Chantelle Sutton
University of Victoria

Speakers use gestures to communicate within a dialogue, not as isolated individuals. We therefore analyzed gestural communication within dyadic dialogues. Specifically, we microanalyzed *grounding* (the sequence of steps by which speaker and addressee ensure their mutual understanding) in a task that elicited abstract deictic gestures. Twenty-two dyads designing a hypothetical floor plan together without writing implements often used gestures to describe these non-existent spaces. We examined the 552 gestures (97% of the database) that conveyed information that was not presented in the accompanying words. A highly reliable series of analyses tracked the immediate responses to these nonredundant speech/gesture combinations. In the vast majority of cases, the addressee's response indicated understanding, and the speaker/gesturer's actions confirmed that this understanding was correct.

1. Studying gestural communication by individuals versus dyads

Laboratory studies of gestural communication usually focus on the speaker and the addressee separately, as encoder or decoder. In encoding studies, the focus is on gesture production in differing conditions (e.g., how visibility influences the speaker's gestures; see review in Bavelas, Gerwing, Sutton, & Prevost 2008, Table 1). Because only the speaker's actions are of interest, the task and the interaction are highly asymmetrical. In these dialogues, the addressee, who may be the experimenter, a confederate, or another participant, often has instructions to respond minimally. Unfortunately, research has shown that constraining the addressee's behaviors may have an unintentional, deleterious effect on the speaker's communicative behaviors (Bavelas, Coates, & Johnson 2000, Beattie & Aboudan 1994).

Decoder studies focus primarily on gesture comprehension (see review in Holler, Shovelton, & Beattie 2009). These designs can be even more removed from dyadic conversation. For example, the participants might watch gestures in brief video

© 2011. John Benjamins Publishing Company
All rights reserved

excerpts, often without conversational context. Again, evidence from other fields suggests that such a design would affect the addressee's ability to understand the gestures. For example, Schober and Clark (1989) found significantly better comprehension by an addressee who was interacting with the speaker than by someone who heard the same information but did not participate in the dialogue. Thus, the encoder and decoder research designs we have been using are not well suited to investigating conversational gestures, which by definition occur within real dialogues.

Recent research has begun to include experiments with two freely interacting participants (e.g., Bangerter 2004; Bavelas, Chovil, Coates, & Roe 1995; Bavelas, Chovil, Lawrie, & Wade 1992; Bavelas et al. 2008; Clark & Krych 2004; Furuyama 2000; Gerwing & Bavelas 2004; Holler & Stevens 2007; Özyürek 2000, 2002). However, the unit of analysis in many of these experiments has remained individual in the sense that the dependent variable was usually a summary of one participant's gestures (e.g., average rate of speaker's gestures). Such measures of aggregated individual actions are useful or even essential for answering certain experimental questions, but they necessarily remove communicative acts from their sequential context, separating one participant's actions from the immediately preceding and succeeding actions of the other person.

In three of the above studies, the dependent variable did reflect the immediate dyadic sequence in which the gestures occurred. Bavelas et al. (1995, Study 2) demonstrated that addressees responded as predicted to the speaker's spontaneous interactive gestures. Furuyama (2000) illustrated how addressees sometimes incorporated the speaker's previous gesture into their own. Clark and Krych (2004) demonstrated how addressees used gestural actions to indicate their state of understanding of the speaker's directions. In each of these three studies, the primary focus was on a gesture in relation to the immediate dyadic context in which it occurred, and the summary data preserved this unit of analysis.

We propose that the participants in a conversation shape their gestures, like their words, to fit a specific, immediate context. Therefore, the ideal design for revealing whether and how conversational gestures communicate would focus on dyadic sequences and would include (a) two or more participants who can interact spontaneously and as themselves; (b) a symmetrical task to which both can contribute; (c) the gestures of both participants; and (d) an analysis of each gesture in the context and interactive sequence within which it occurred. In pursuit of this ideal, the present study obtained moment-by-moment dyadic evidence of gestural communication using a design that included two real participants, without constraints on their interaction, designing a floor plan together. The gestures could be from either participant, and our analysis of grounding sequences included the responses of both of them.

2. Grounding

Fundamental to Clark's (1996) collaborative model of language use is *grounding* (Clark & Schaefer 1989), a moment-by-moment process by which the participants establish

© 2011. John Benjamins Publishing Company
All rights reserved

that they understand each other well enough for current purposes. Grounding is an intrinsically mutual responsibility, not an individual process: "Speakers and their addressees go beyond autonomous actions and collaborate with each other, moment by moment, to try to ensure that what is said is also understood" (Schober & Clark 1989: 211). "Moment by moment" means that grounding is a micro-process that is constantly occurring, usually in the background of the dialogue and not just in conclusion.

Our preferred description of a *grounding sequence* involves a rapid three-step interchange between the participants: The person who is speaking at the moment *presents* some information, the addressee *responds* with an indication or display of understanding (or not), and then the speaker *acknowledges* this response by indicating that the addressee's understanding was correct (or not). These steps can involve words, gestures, nodding, gaze, or other actions, singly or in combination.

In the following examples from our floor-plan data, underlined words indicate the location of a gesture. Also, throughout this chapter, we will distinguish between the participants by arbitrarily treating the speaker/gesturer of the moment as female and the addressee at that moment as male.

(1) The speaker was describing their plan, starting at the entrance to the apartment:
Speaker: So we could have, like, <u>you come in</u>.
Addressee: Yeah.
Speaker: There's <u>a kitchen</u> ...

While saying "<u>you come in</u>," the speaker gestured the location of the entrance by placing her two index fingers together on the table. The addressee indicated explicitly that he understood the location by saying "Yeah." Then the speaker/gesturer located "<u>a kitchen</u>" by placing her left hand slightly to the left of where she had placed the entrance. Notice that, instead of explicitly acknowledging the addressee's understanding, the speaker/gesturer presupposed it by continuing her tour of the floor plan.

Addressees also use continuation as a way of indicating understanding:

(2) The participants were reviewing their plan, and the speaker had just used gestures to place the two bedrooms on either side of a hallway.
Speaker: ... and then <u>a bathroom</u>
Addressee: bathroom at the end
Speaker: [nods]

As the speaker said "and then <u>a bathroom</u>," she pointed to a spot at the end of where she had previously placed the hallway. The addressee immediately displayed his understanding by saying "bathroom" simultaneously and then finishing her sentence by naming the location that the speaker had only gestured ("at the end"). The speaker's nod explicitly acknowledged that the addressee had understood correctly.

Recall that the standard for grounding is "well enough for current purposes" (Clark 1996: 221), so the participants may also rely on implicit indications of

© 2011. John Benjamins Publishing Company
All rights reserved

understanding. Indeed, conversation would sink under its own weight if every step of every grounding sequence were explicit. Instead, participants often minimize their joint effort by more economical implicit responses, as shown in the next example. Note that there were two presentations in this example, and the grounding was entirely implicit in the first one:

(3) Speaker: In my mind, $_1$<u>the bedrooms</u>... $_2$<u>are on this side</u>.
Addressee: [nodding] Ohhh-kay!
Speaker: Yeah.

The speaker/gesturer began the first sequence with the words "In my mind, <u>the bedrooms</u>" as she placed her hand to show the location of one of the bedrooms. She then paused briefly, and the addressee continued to watch her gestures (implicitly indicating understanding). The speaker then said "<u>are on this side</u>" while moving her hand to a location further beyond, where the other bedroom would be. This second presentation of new information served two functions: It presupposed the addressee's understanding of her first gesture, thereby implicitly acknowledging it and ending that grounding sequence, and it presented further new information, initiating a new sequence. This time, the addressee indicated his understanding explicitly (with "Ohhh-kay!" and a big nod), and the speaker/gesturer's acknowledgment was also explicit ("Yeah").

A grounding analysis can also identify points at which mutual understanding does not occur. At each step, either participant can initiate a clarification or repair. That is, the addressee can ask for clarification from the speaker/gesturer. Or the speaker/gesturer can detect that the addressee's understanding is wrong and correct it.

In sum, grounding sequences are an observable, intrinsically dyadic process, focused precisely on the establishment of mutual understanding. They are thus well-suited to examining the communicative value of gestures for interlocutors. Our analysis focused on the grounding process initiated by presentations of nonredundant speech/gesture combinations (i.e., ones where the gesture conveyed information that was otherwise missing from the words), then examined the addressee's immediate response, and then the speaker/gesturer's acknowledgment. We propose that a successful grounding sequence after a nonredundant speech/gesture combination provides observable, local evidence that the participants used these gestures to communicate and mutually considered the gestural information to be part of their accumulating common ground.

3. Abstract deictics

The task used here evoked a different kind of gesture than in many previous experiments, namely, gestures depicting something that does not exist. The participants sat across a bare table and designed a floor plan for a student apartment. As they talked, all of them spontaneously "drew" their plans on the table with their gestures, creating

© 2011. John Benjamins Publishing Company
All rights reserved

and pointing to hypothetical spatial relationships that had no concrete referent. These gestures were *abstract deictics* (e.g., McNeill, Cassell, & Levy 1993), which are a special kind of pointing. As Kendon (2004) explained, most pointing gestures indicate a space or location that is currently visible or a direction toward a real location that is not yet visible. In contrast, abstract deictics actually create spaces and refer to locations that do not physically exist. Our participants' gestures did not represent any existing space; they depended entirely on the participants' shared understanding of their words and gestures. We expected that, even in these cases, the participants would readily show that they understood each other.

4. Research design and procedures

4.1 Task and hypotheses

Each dyad designed its own layout for a two-bedroom student apartment on the table between them.[1] There were no assigned roles; both participants could contribute to the design of the plan as they wished. We emphasized the goal of mutual understanding by advising them that when they were finished, they would each have to draw the agreed-upon plan independently.

4.2 Method and procedure

A total of 44 University of Victoria students formed 22 dyads (12 female/female, 1 male/male, and 9 female/male). All participants spoke English fluently, were unacquainted, and knew they would be videotaped. In return for participating, they received course bonus credits.

Recording equipment in our Human Interaction Lab included a remotely controlled Panasonic WD-D5000 color camera with a wide-angle lens and a Soundgrabber II omni-directional microphone. We digitized the analog video into AVI format using Broadway (www.b-way.com) and analyzed it with Broadway on an 18-inch ViewSonic GS790 color monitor.

After the participants read and signed a consent form, they had a few minutes to get acquainted with one another. They then did two or three unrelated tasks, including the primary one: The experimenter asked them to design a floor plan for a two-bedroom apartment appropriate for University Student Housing. The floor plan should include (but not necessarily be limited to) the bedrooms, a bathroom, a living room, and a kitchen. The experimenter emphasized that the layout of the apartment was

1. We varied the width of the table the dyad worked on. As predicted, the wider space led participants to move their gestures forward, toward their partner. Because there were no other significant differences, we will not include this variable in the rest of the chapter.

© 2011. John Benjamins Publishing Company
All rights reserved

most important, not the dimensions of the rooms or where the furniture went. She also informed them that, later, they would each have to draw the floor plan separately. After answering questions, the experimenter left the participants to design their plan. When they were done, she returned and re-seated them on either side of a partition to make their individual drawings of the plan.

Afterward, the experimenter explained the purpose of the study, answered questions, and gave them a written summary. Finally, they watched their videotape, and each indicated on a permission form whether and how we could use their videotape (e.g., to be viewed only for research, shown to professional audiences, etc.).

5. Analysis and results

5.1 The data set

The purpose of the analysis was to examine grounding sequences that began whenever either of the participants used a nonredundant combination of speech and gesture to present new information about the location of a room or rooms in their floor plan. We limited the potential data to about two minutes of each dyad's discussion of their final floor plan, excluding initial discussions of possible criteria and preliminary layouts. When a discussion of the final plan was substantially longer than two minutes, we analyzed only the first and last minute. During these two minutes, the mean proportion of time spent gesturing was .82 ($SD = .11$).

Within this data set, independent analysts located all gestures that depicted an identifiable room. They excluded gestures that were not about the floor plan; gestures that did not locate an identifiable room within it; gestures that were not analyzable after repeated viewing; and adaptors. They included gestures by the addressee of the moment only if the response added new verbal or gestural information, initiating a new, overlapping grounding sequence. Note that our focus was not on individual gestures but on the presentation of information about rooms in the plan, which could include one or more rapidly contiguous gestures. The inter-analyst reliability for the above decisions ranged from 80% to 97%.

The final data set was 571speech/gesture combinations that depicted identifiable rooms in the floor plan.

5.2 Identifying nonredundant speech/gesture combinations

We focused on *nonredundant* gestures, which contributed information that was missing from the words. Nonredundant gestures required that the addressee apprehend and integrate information from both speech and gesture. A typical nonredundant gesture was

(4) Speaker: Let's say we have the door here.

© 2011. John Benjamins Publishing Company
All rights reserved

As she said "door here," the speaker/gesturer traced a line about an inch wide on the table. It was only her gesture that showed precisely where "here" was. Therefore, the gesture was nonredundant with the words. Notice that all of the gestures in Examples 1, 2, and 3 above were also nonredundant.

In contrast, *redundant* gestures conveyed no additional information beyond the words; for example,

(5) Speaker: so we put the bedrooms on the right side and the bathrooms on the left, is that right?

The speaker/gesturer first used her right hand to make a vague pointing gesture to her right; then, she used her left hand to make a similar gesture to her left. Both her words and her gestures depicted "right" then "left," with no additional or more specific information in the gestures, which were therefore redundant with her words. See Gerwing and Allison (2009) for a more detailed explanation of this and other redundancy analyses.

Reliability for redundancy versus nonredundancy across all groups and all gestures was 96.5%.

5.2.1 *Redundancy results*

Redundancy between gestures and words was rare; 552 of the 571 speech/gesture combinations analyzed included gestures that were *not* redundant with the words (mean proportion = .97; *SD* = .05). As illustrated in examples 1 to 4, the gestural information was usually essential to their task (e.g., the location of the rooms).

5.3 Grounding sequences

A grounding sequence consisted of the *presentation* of one of the above 552 nonredundant speech/gesture combinations, the *addressee's response,* and any *acknowledgement* by the speaker/gesturer. Figure 1 depicts the overall analysis.

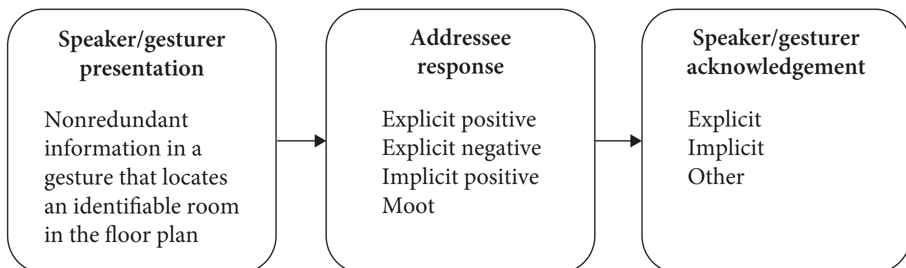

Figure 1. Schematic figure of the three stages of analysis.

© 2011. John Benjamins Publishing Company
All rights reserved

5.3.1 *Addressee responses to nonredundant gestures*

Immediately following each of the speaker/gesturer presentations, addressees could respond by indicating whether they understood (or not). As described below, their response could be explicit or implicit. It could also be positive (indicating understanding), negative (indicating lack of understanding), or moot (indeterminate).

5.3.1.1 *Explicit versus implicit addressee responses.* An *explicit addressee response* was one that provided decisive feedback to the speaker/gesturer about whether the addressee had or had not understood the nonredundant speech/gesture combination. Examples included saying "yeah," finishing the speaker/gesturer's sentence, nodding, gesturing the same room, or alternatively, asking a question. However, in the second-by-second tempo of a spontaneous dialogue, it would be inefficient for addressees to respond explicitly to every presentation that they had understood; the participants sometimes rely on implicit indications. *Implicit addressee responses* did not provide overt evidence about the addressee's state of understanding. The addressee simply continued to pay attention and allowed the speaker/gesturer to go on, or the addressee took an action that implicitly built on the speaker/gesturer's presentation without any overt expression of understanding. Two analysts examined what the addressee did immediately following the speaker/gesturer's presentation and decided whether the addressee contributed an explicit or implicit response. Their reliability on a randomly selected set of 19 groups was 89%.

5.3.1.2 *Explicit positive versus explicit negative responses.* Explicit addressee responses could be positive, indicating understanding, or negative, indicating not understanding or requesting clarification. Typical *explicit positive responses* were "yeah," nodding, or gesturing the same room. In an *explicit negative response*, the addressee typically asked for clarification about the relative location of rooms. Based on a randomly selected 20 groups, inter-analyst reliability for distinguishing whether an explicit addressee response was positive or negative was 96%.

5.3.1.3 *Implicit positive versus moot.* Recall that the standard for grounding is "well enough for current purposes" (Clark 1996: 221), so it is efficient for the participants to use some implicit indications of understanding. However, it is more difficult for analysts, who are outside the dialogue, to judge when an implicit response is clearly negative. Therefore, in our analysis, implicit addressee responses could be either positive or moot. *Implicit positive responses* occurred when the addressee did not overtly indicate a lack of understanding. He either continued to watch the speaker/gesturer or said something that built on a presupposed understanding. The remaining cases were *moot*; the addressee was either looking away from the speaker/gesturer or said something unrelated to the previous presentation of information, possibly overlooking or ignoring the speaker's contribution. We deemed these responses to be moot because they were not even implicitly positive. First, two analysts examined all

© 2011. John Benjamins Publishing Company
All rights reserved

implicit responses and, together, differentiated between positive and moot addressee responses. Then a third analyst did the same analysis independently for six of the 22 dyads. Reliability between the first pair and the third analyst was 91%.

5.2.1.4 *Results for addressee responses.* The results provided strong, moment-by-moment dyadic evidence that addressees understood presentations with nonredundant gestures. The vast proportion of their responses were positive (M = .955, SD = .052) rather than negative or moot (M = .045, SD = .051). A paired-sample t-test indicated that these mean proportions were significantly different ($t_{(21)}$ = 41.7, p < .001). Moreover, even though the dialogues often had rapid or even overlapping exchanges, the addressees were more likely to provide explicit feedback (M = .617, SD = .140) than only implicit feedback (M = .382, SD = .139); again, this difference was significant ($t_{(21)}$ = 3.95, p = .001). As suggested by these mean proportions, the addressees' positive responses were more often explicit (M = .589, SD = .152) than implicit (M = .366, SD = .143); $t_{(21)}$ = 3.60, p < .01. That is, addressees were significantly more likely to provide overt evidence that that they had understood the speaker/gesturer's presentation than inferential evidence. It is noteworthy that explicit negative responses (indicating that the addressee had *not* understood) were extremely rare (M = .029, SD = .036). All of these 17 instances were questions. Our impression was that, in about half of these cases, the addressee was seeking to clarify a genuine misunderstanding. In the remaining cases, the addressee may have understood and was asking a question as a polite way of disagreeing (e.g., "Oh so you walk through the kitchen into the living room?"). Finally, the mean proportion of implicit addressee responses that were moot was also very small (M = .016, SD = .035).

5.3.2 *Acknowledgment by the speaker/gesturer*

In a fully explicit grounding sequence, the speaker/gesturer would acknowledge (or correct) the addressee's indication of understanding. However, constantly stopping the flow of content to acknowledge the correctness of the addressee's understanding would be quite inefficient, violating the principle of *least joint effort* (e.g., Clark 1996, Clark & Krych 2004, Clark & Schaeffer 1989). Indeed, this third step in the grounding sequence does not appear in many versions of the theory (e.g., Clark, 1996). Most versions treat the speaker's confirmation of understanding as the default response, which would therefore be implicit. We tested this assumption by examining what the speakers in our task actually did to close each grounding sequence.

5.3.2.1 *Speaker/gesturer's acknowledgment.* We analyzed three possible responses. An *explicit acknowledgment* was analogous to an explicit addressee response; the speaker/gesturer responded overtly, e.g., saying "right" or "OK," nodding, finishing the addressee's sentence, or repeating the addressee's exact word(s). An *implicit acknowledgment* occurred when the speaker/gesturer's response presupposed that the addressee had understood so far. For example, when the speaker/gesturer simply went on to finish

© 2011. John Benjamins Publishing Company
All rights reserved

what she had been saying before the addressee had responded, or she clarified the information in the addressee's response without overt acknowledgment (e.g., did not say "yeah"), or she continued by presenting new information. There were also some *other* responses, such as when the speaker/gesturer said or did something unrelated to the addressee's response or the addressee took up the turn before the speaker/gesturer could continue. Two analysts made these decisions independently for19 of the 22 groups, with 87% agreement for all gesture sequences within those groups.

5.3.2.2 *Results for acknowledgments.* The speaker/gesturers' acknowledgment of addressees' understanding was seldom explicit (M = .151, SD = .081). Instead, they usually acknowledged implicitly, such as by moving on to new information (M = .767, SD = .078). There were few "other" responses (M = .081, SD = .073), which strongly suggests that both participants were completing each grounding sequence (albeit implicitly) rather than interrupting it with other actions.

Recall that the 17 instances of explicit negative addressee responses were questions. The speaker/gesturer's response in 15 of these instances was to answer the question or otherwise clarify what she had presented, usually within a few seconds. That is, the speaker acknowledged the state of the addressee's understanding by providing the required information.

5.3.3 *Results for the grounding sequences*

A grounding sequence is a sequence of contingent actions, and Table 1 shows the proportional relationships between the addressees' and the speaker/gesturers' responses. In the most frequent pattern (42% of the sequences), the addressee indicated his understanding explicitly (e.g., saying "yeah" or repeating the words), then the speaker/gesturer followed up implicitly (e.g., continuing on to new information).

In the next most common pattern (36% of the sequences), the addressee responded implicitly (e.g., simply continued to pay attention), and the speaker/gesturer also carried on implicitly. It is noteworthy that in these cases, the speaker/gesturer did not explicitly check on her addressee's level of understanding. The speaker/gesturer seemed to have acted on the default assumption that the information in her speech/gesture

Table 1. Sequential proportions of addressee responses and speaker follow-up responses

Addressee response	Speaker follow-up response	M (SD)
Explicit	Explicit	.14 (.08)
	Implicit	.42 (.13)
	Other	.05 (.06)
Implicit	Explicit	.01 (.02)
	Implicit	.36 (.13)
	Other	.02 (.03)

© 2011. John Benjamins Publishing Company
All rights reserved

combination was successfully grounded *unless* the addressee explicitly revealed that it was not. Much less frequently (14% of sequences), both participants grounded explicitly. All of the remaining combinations were rare.

6. Summary

Ultimately, it is the participants themselves who determine the communicative value of their gestures. We tested their mutual understanding using a microanalysis of grounding sequences after each nonredundant speech/gesture combination that located a room or rooms in their proposed floor plan. Gestures are essential in such spatial tasks, and virtually all of the gestures they used carried information that was not in their words. The addressee had to understand both the words and gestures together.

Mutual understanding was potentially even more difficult in this task because the gestures lacked any external anchor or referent. There were no real objects or spaces to point at or manipulate. The dyad had to co-construct and sustain the invisible floor plan with their words and abstract deictic gestures. In spite of the difficulty of their task and the speed of spontaneous dialogue, only 4.5% of the addressees' 552 responses indicated that they had not understood the information that the speaker/gesturer had presented.

The results suggest that this method would be useful both for looking even more closely at how dyads understand each other's gestures and for examining the process in other situations. Grounding is an "opportunistic" process (Schober & Clark 1989) in which the participants seize on whatever works, and solutions to grounding in other contexts could not fail to be interesting.

References

Bangerter, A. 2004. "Using pointing and describing to achieve joint focus of attention in dialogue." *Psychological Science* 15 : 415–419.

Bavelas, J., Chovil, N., Coates, L. and Roe, L. 1995. "Gestures specialized for dialogue." *Personality and Social Psychology Bulletin* 21 (4): 394–405.

Bavelas, J., Chovil, N., Lawrie, D. A. and Wade, A. 1992. "Interactive gestures." *Discourse Processes* 15: 469–489.

Bavelas, J., Coates, L. and Johnson, T. 2000. "Listeners as co-narrators." *Journal of Personality and Social Psychology* 79 (6): 941–952.

Bavelas, J., Gerwing, J., Sutton, C. and Prevost, D. 2008. "Gesturing on the telephone: Independent effects of dialogue and visibility." *Journal of Memory and Language* 58: 495–520.

Beattie, G. and Aboudan, R. 1994. "Gestures, pauses and speech: An experimental investigation of the effects of changing social context on their precise temporal relationships." *Semiotica* 99: 239–272.

Clark, H. H. 1996. *Using Language*. Cambridge, UK: Cambridge University Press.

© 2011. John Benjamins Publishing Company
All rights reserved

Clark, H. H. and Krych, M. A. 2004. "Speaking while monitoring addressees for understanding." *Journal of Memory and Language* 50: 62–81.

Clark, H. H. and Schaefer, E. F. 1989. "Contributing to discourse." *Cognitive Science* 13: 259–294.

Furuyama, N. 2000. "Gestural interaction between the instructor and the learner in origami instruction." In *Language and Gesture*, D. McNeill (ed), 99–117. Cambridge: Cambridge University Press.

Gerwing, J. and Allison, M. 2009. "The relationship between verbal and gestural contributions in conversation: A comparison of three methods." *Gesture* 9 (3): 312–336.

Gerwing, J. and Bavelas, J. 2004. "Linguistic influences on gesture's form." *Gesture* 4 (2): 157–195.

Holler, J., Shovelton, H. and Beattie, G. 2009. "Do iconic hand gestures really contribute to the communication of semantic information in a face-to-face context?" *Journal of Nonverbal Behavior* 33: 73–88.

Holler, J. and Stevens, R. 2007. "The effect of common ground on how speakers use gesture and speech to represent size information." *Journal of Language and Social Psychology* 26 (1): 4–27.

Kendon, A. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

McNeill, D., Cassell, J. and Levy, E. T. 1993. "Abstract deixis." *Semiotica* 95: 5–19.

Özyürek, A. 2000. "The influence of addressee location on spatial language and representational gestures of direction." In *Language and Gesture*, D. McNeill (ed), 64–83. Cambridge: Cambridge University Press.

Özyürek, A. 2002. "Do speakers design their cospeech gestures for their addressees? The effects of addressee location on representational gestures." *Journal of Memory and Language*, 46: 688–704.

Schober, M. F. and Clark, H. H. 1989. "Understanding by addressees and overhearers." *Cognitive Psychology* 21: 211–232.

© 2011. John Benjamins Publishing Company
All rights reserved