

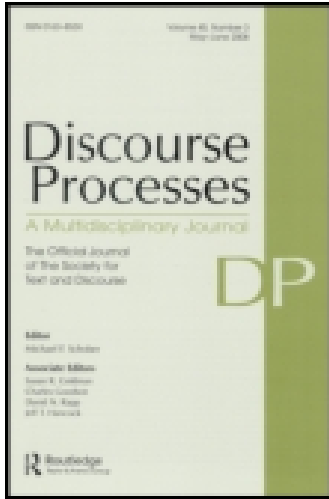
This article was downloaded by: [University of Victoria]

On: 10 July 2015, At: 12:18

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954

Registered office: 5 Howick Place, London, SW1P 1WG



Discourse Processes

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/hdsp20>

Effect of Dialogue on Demonstrations: Direct Quotations, Facial Portrayals, Hand Gestures, and Figurative References

Janet Bavelas^a, Jennifer Gerwing^a & Sara Healing^a

^a Department of Psychology, University of Victoria, Victoria, British Columbia, Canada

Accepted author version posted online: 24 Jan 2014. Published online: 13 Oct 2014.



CrossMark

[Click for updates](#)

To cite this article: Janet Bavelas, Jennifer Gerwing & Sara Healing (2014) Effect of Dialogue on Demonstrations: Direct Quotations, Facial Portrayals, Hand Gestures, and Figurative References, *Discourse Processes*, 51:8, 619-655, DOI: [10.1080/0163853X.2014.883730](https://doi.org/10.1080/0163853X.2014.883730)

To link to this article: <http://dx.doi.org/10.1080/0163853X.2014.883730>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages,

and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Effect of Dialogue on Demonstrations: Direct Quotations, Facial Portrayals, Hand Gestures, and Figurative References

Janet Bavelas, Jennifer Gerwing, and Sara Healing

Department of Psychology

University of Victoria, Victoria, British Columbia, Canada

Demonstrations (e.g., direct quotations, conversational facial portrayals, conversational hand gestures, and figurative references) lack conventional meanings, relying instead on a resemblance to their referent. Two experiments tested our theory that demonstrations are a class of communicative acts that speakers are more likely to use in dialogue than in monologue. We compared speakers' rates of demonstrations in face-to-face dialogues, telephone dialogues, and monologues into a handheld microphone or recorder. Experiment 1 confirmed that the proportions of speakers' direct quotations and facial portrayals were (1) significantly higher in the two dialogue conditions than in the monologue condition and (2) not significantly different in the two dialogue conditions. Experiment 2 found the same patterns for the rates of figurative references and hand gestures. In both experiments, regressions confirmed that the increase in demonstrations in dialogues was independent of any effect of visibility.

Jennifer Gerwing is now also a postdoctoral fellow at the Health Services Research Unit, Akershus University Hospital, Oslo, Norway.

Correspondence concerning this article should be addressed to Janet Bavelas, Department of Psychology, University of Victoria, P.O. Box 1700 STN CSC, Victoria, BC V8W 2Y2, Canada. E-mail: bavelas@uvic.ca

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hdsp.

INTRODUCTION

Clark and Gerrig's (1990) *demonstrations* are a coherent class of communicative acts that are different from *descriptions*. Whereas descriptions have arbitrary or conventional meanings, demonstrations present a selective depiction of their referent and rely on this resemblance for their meaning. The present work makes three proposals: First, according to Clark and Gerrig's theory, direct quotations, conversational facial portrayals, conversational hand gestures, and figurative language are all demonstrations. Second, as Bavelas, Gerwing, Sutton, and Prevost (2008, pp. 517–519) postulated, demonstrations as a class share another common feature, namely, that speakers use them at higher rates in dialogue than in monologue. Third, this effect of dialogue versus monologue is distinct from any effect of mutual visibility.

After illustrating the common features of the four demonstrations, we present two experiments testing the dialogue hypothesis: Speakers in a dialogue with an addressee (whether mutually visible or not) would make all four demonstrations at a significantly higher rate than speakers alone in a monologue. The concluding discussion outlines possible explanations for the effect of dialogue on demonstrations, but the central purpose of the present experiments is to test the generality of our theory by confirming that the effect of dialogue holds over a wide range of demonstrations.

What Are Demonstrations?

Clark and Gerrig (1990; see also Clark, 1996, Chapter 6) transformed Peirce's (e.g., 1960) trichotomy of *symbols*, *indices*, and *icons* into three methods of signaling that interlocutors can use in a dialogue: *describing*, *indicating*, and *demonstrating*. The following example illustrates and contrasts these three methods. Imagine a shopper who approaches the salesperson in a furniture store. The shopper has three different options for how to tell him what she wants:

- *Description*: "I'm looking for *an oval dinner table*," using verbal symbols (i.e., words) in their conventional, arbitrarily established meanings.
- *Indication*: "I'm looking for a dinner table *like the one over there*" while pointing at an oval table nearby. Notice that, to use this option, both the shopper and the salesperson must be able to see an oval-shaped table; that is, the referent must be present.
- *Demonstration*: "I'm looking for a dinner table *that's shaped like this*" while sketching a symmetrical oval shape (e.g., starting with her fingers tips together at the top, then moving apart, rounding down, and coming together again at the bottom). The gesture presents an improvised (i.e., not standardized) image of an oval shape, which does not require its referent to be present.

The gesture in the third example illustrates four distinguishing characteristics of demonstrations. First, demonstrations create a *version of their referent*; that is, they “work by enabling others to experience what it is like to perceive the things depicted” (Clark & Gerrig, 1990, p. 765). Second, a demonstration must therefore *resemble* its referent in some way. In the shopper’s gesture, the salesperson could observe directly the oval shape that the shopper is looking for. Third, although a demonstration resembles its referent, it is also the speaker’s *transformation* of the literal or actual properties of the referent (e.g., a table top is a horizontal surface, but she draws it on the vertical plane in the gesture space in front of her). Finally, this transformation is *selective* (Clark & Gerrig, 1990, pp. 767–769). The shopper depicts only the oval shape of a hypothetical table surface and not its size, style, or other design features.

In a dialogue, the interlocutors have available at least four different kinds of demonstrations: direct quotations, facial portrayals, hand gestures, and figurative references. The next sections describe¹ and illustrate each kind of demonstration in more detail.

Direct quotations. Clark and Gerrig (1990) introduced their theory of demonstrations with *direct quotations*, in which the speaker creates a version of what someone said or might have said. Although a direct quotation resembles its referent, speakers are not claiming to be replicating the original utterance exactly. Instead, speakers are providing a version of what was said. For example, in Figure 1, a speaker in a face-to-face dialogue was retelling a scene from the movie *Shrek 2*. In the scene, Shrek picks up a cat by the back of the neck and holds him close to his face while the cat begs for his life. In frame 3, the speaker quoted the cat’s plea in her own words, “And then the cat was like, ‘Oh, I’m so sorry, I’m so sorry.’” In the movie scene, the cat was begging for mercy and explaining his actions at some length but did not actually say he was sorry. Notice that in frames 1 and 2, the speaker referred to the cat in third person (“him,” “his,” and “the cat”) but switched to first person, using “I” for the cat when she was quoting him.

Conversational facial portrayals. Kraut and Johnston (1979) as well as Ekman (1997) distinguished between facial expressions of emotion and the more varied social functions of faces that occur in interaction. In the first systematic study of faces in dialogue, Chovil (1989, 1991/1992) found that interlocutors’ faces in dialogue displayed a virtually unlimited variety of rapid (.5 s or less)

¹To prevent confusion, we will continue to use *describe* and *description* in their ordinary meanings (i.e., to convey, relate, give the details of, etc.) and will note explicitly when using *description* in Clark and Gerrig’s (1990) sense.



FIGURE 1 An example of a participant making three kinds of demonstrations while retelling a scene from *Shrek 2*: a hand gesture in frame 2, facial portrayals in frames 2 and 3, and a direct quotation in frame 3.

facial actions directly related to speech. Bavelas and Chovil (1997) proposed that these conversational facial actions are demonstrations: The speaker's face creates selective images that are related to the topic of the dialogue at that moment rather than being literal expressions of what the speaker is feeling. This article focuses on a specific kind of facial gesture, speakers' *facial portrayals* of someone other than themselves. In Figure 1, the speaker demonstrated her version of the faces of two characters in her narrative: In frame 2, exactly as she finished her sentence with "by the nape of his neck," her face portrayed Shrek's somewhat malevolent pleasure as he held the cat at his mercy. In frame 3, while saying "Oh I'm so," she portrayed the cat's worried expression. These facial demonstrations were very brief; the first one lasted .41 s and the second one lasted .37 s.

Conversational hand gestures. In a natural dialogue, speakers spontaneously improvise *conversational* (or *co-speech*) *gestures*, which are hand movements that convey meanings closely related to what the gesturer is saying at that precise moment. Conversational hand gestures have all the features of demonstrations: The images they present are transformations of a referent, portraying only selected aspects of it. In frame 2 of Figure 1, exactly while saying "the nape of his neck," the speaker used a hand gesture to present the addressee with an image of Shrek holding the cat between his fingers and thumb, at his eye level. She pinched her first two fingers to her thumb while raising her hand as if holding the suspended cat in front of her face. The speaker's hand represented Shrek's hand, and the proximity of her hand to her face mirrored the proximity between Shrek's hand and face in the movie scene. Her gesture was not an exact replication of Shrek's actions; it was her selective demonstration of those actions.

Figurative references. According to both Peirce (1940, p. 105) and Clark (1996, p. 157), metaphors are also icons or demonstrations, and, like quotations, they create images with words. For example, Peirce (1960) pointed out that such an icon "may require its interpretation to involve the calling up of an image" (p. 360).

We are including here a broader class than formally structured metaphors, namely, *figurative references*, which demonstrate a property of the referent by using the features or characteristics of something else. When describing a scene in *Shrek 2* in which the cat was silently pleading with Shrek, some speakers said that the cat had (or made) “puppy-dog eyes,” which calls up an image of very large, round eyes demonstrating an appealing helplessness (vs. shock or fear).

Dialogue–Monologue as a Variable

Bavelas et al. (2008) found that both hand gestures and figurative references occurred at significantly higher rates in dialogue than in monologue, which led to the present hypothesis that dialogue might also increase other demonstrations, such as direct quotations and facial portrayals. To pursue this hypothesis, it is necessary to examine more closely the characteristics of dialogue and monologue.

In a face-to-face dialogue, interlocutors are (1) mutually visible and (2) taking part in a social interaction. Definitions in the literature often imply that these two characteristics are indivisible. For example, Luckmann (1990) stressed that “dialogue is always concretely actualized as part of face-to-face social interaction” (p. 52). Marková (1990) defined dialogue as “symbolic face-to-face oral and gestural communication” but added that “the special characteristics of dialogue . . . are a result of interaction” (p. 6). More recently, however, Linell (2009) pointed out that although dialogue is “face-to-face interaction through talk” (p. 4), it can also be mediated (e.g., via the telephone or email). That is, whether the interlocutors can interact in a dialogue is distinct from whether they can see each other.

Clark’s 10 features of face-to-face dialogue (1996, pp. 9–10) include both interaction and mutual visibility. Three of these features define interaction in a free dialogue: (1) Both interlocutors are *acting as themselves* (e.g., one is not a confederate); (2) their actions are *self-determined* (i.e., not scripted; they decide for themselves what to do and when); and (3) they are acting *extemporaneously, in real time* (i.e., improvising in the moment, not rehearsed). Together, these criteria suggest that creating a dialogue is not as simple as providing the speaker with an addressee. Table 1 draws on existing definitions and experimental procedures to suggest a continuum of possible contexts from a free dialogue to an extreme monologue.

The dialogue conditions in the present experiments corresponded to the first context in Table 1, with both speaker and addressee interacting freely and spontaneously within their given task. The monologue condition corresponded to the seventh context: The speaker was alone in the room, describing a stimulus to the best of his or her ability into a handheld recorder or microphone. This condition fits the broader tenets of monologism (e.g., Linell, 2009; Marková, 1990), which focuses on the abstract language system and individual language

TABLE 1
A Continuum of Contexts From Dialogue to Monologue

-
1. Both the speaker and the addressee are interacting freely, that is, extemporaneously, determining their own actions and taking actions as themselves (Clark, 1996, p. 10).
 2. The speaker is talking to a distracted and therefore less responsive addressee (e.g., Bavelas, Coates, & Johnson, 2000).
 3. The speaker is talking to an addressee who is not acting as him- or herself and is at least partially scripted (e.g., a confederate, the experimenter, or a participant whose actions are constrained by instructions—a common context in social psychology).
 4. The speaker has an audience of listeners, but they are not individual addressees. For example, “One person [is speaking to an audience] with little or no opportunity for interruption or turns by members of the audience” (Clark, 1996, p. 4) or “speech . . . by a single person, as in a lecture” (Crystal, 2001, pp. 220–221).
 5. The speaker is alone but practicing what he or she will say later, such as “preparing and listening to speeches” (Garrod & Pickering, 2004, p. 8). For example, in the “alone” condition in Cohen’s (1977) gesture experiment, the speaker was rehearsing his descriptions for later presentation to the experimenter.
 6. The speaker is talking to an imagined addressee, who does not exist (e.g., Fridlund, 1991; Fridlund et al., 1990).
 7. The speaker is alone and is focusing on his or her own ability to describe or respond to something, without reference to an addressee or audience. This context is not uncommon in psychology experiments.
 8. The speaker is alone and believes that what he or she says will never be seen or heard by anyone else. This situation is unlikely to be achievable within an experiment.
-

users rather than on how users interact when talking with each other. The contrast between our dialogue and monologue conditions corresponds to Glucksberg and Krauss’s (1967) distinction between *social* and *nonsocial speech*. Yule (1997, pp. 2–3) defined the difference as follows:

[social speech is] produced specifically to take account of some other (the current listener) and [is] responsive to what the other does, knows, and says. Non-social speech is . . . produced as an expression of the perspective of self.

These two extremes provide a logical first test of whether dialogue and monologue affect demonstrations differently. If there were no differences between a free dialogue and this extreme monologue condition, then the hypothesized dialogue effect could be rejected out of hand.

Controlling for Mutual Visibility

If one accepts that mutual visibility and dialogic interaction are distinct variables, then visibility would be a fatal confound in any comparison between a face-to-face dialogue and a monologue by one person who is alone. It is necessary to have

a second dialogue condition in which the interlocutors cannot see each other (a not-visible condition). If speakers in both dialogue conditions produce higher rates of demonstrations than speakers in the monologue condition, then this effect could not be due to differences in visibility.

We chose a telephone dialogue for the not-visible condition because it is similar to face-to-face dialogue in two important respects. First, talking on the telephone is as familiar as talking in a face-to-face dialogue and certainly more familiar than other alternatives, such as talking through an intercom or partition. Second, Chovil (1991) found that dialogue on the telephone, compared with dialogue through a partition, was ranked as significantly higher in *sociality*, that is, “how close people would feel in the situation and how easily the people would find it to converse with each other” (p. 149). Because the sociality of a dialogue is central to our hypothesis, the telephone is a more appropriate choice.

It must be noted that holding a telephone may at first appear to present a confound for the production of hand gestures in particular. However, Bavelas et al.’s (2008) finding of no significant difference in the overall gesture rate in face-to-face versus telephone conditions is identical to the results of five similar experiments that compared face-to-face versus partition conditions (Bavelas, Chovil, Lawrie, & Wade, 1992, Exp. 2; de Ruiter, Bangertter, & Dings, 2012²; Holler, Tutton, & Wilkin, 2011; Pine, Burney, & Fletcher, 2010; Rimé, 1982). This unusual number of replications strongly suggests that holding a phone does not affect overall gesture rate. However, to eliminate any difference between the two not-visible conditions, the speakers in the monologue condition also held something, namely, they spoke into a handheld recorder or microphone.

Summary: Design and Predictions

Three experimental conditions isolated the effects of dialogue and visibility:

- A face-to-face dialogue (social interaction plus mutual visibility)
- A telephone dialogue (social interaction without mutual visibility)
- A monologue recorded into a microphone (neither social interaction nor mutual visibility)

Experiment 1 focused on the speakers’ rates of direct quotations and facial portrayals, and Experiment 2 focused on the rates of hand gestures and figurative

²De Ruiter et al. (2012) focused on more specific effects of visibility and did not report results for the overall gesture rate. However, J. P. de Ruiter (personal communication, July 13, 2012) kindly provided these results, which confirmed there was no significant visibility effect on the mean rate of iconic gestures per 100 words.

references. Our hypothesis was that there should be a significantly higher rate of all four types of demonstrations in dialogue than in monologue, even after controlling for mutual visibility.

EXPERIMENT 1: DIRECT QUOTATIONS AND FACIAL PORTRAYALS

In telling stories, whether real or fictitious, speakers often briefly portray the story's characters. One speaker might demonstrate what a character said, speaking for the moment as if he himself were the character (a *direct quotation*). Another might use her face to demonstrate how a character looked or reacted, for example, wrinkling her nose to demonstrate a character's expressed disgust (a *facial portrayal*). Both direct quotations and facial portrayals are examples of what McNeill (1992, pp. 118–119) described as taking a *character* rather than *observer* viewpoint. Therefore, we chose a stimulus with characters who provided both quotable dialogue and memorable facial actions—excerpts from the movie *Shrek 2*. In the two dialogue conditions, the speaker and addressee interacted spontaneously and extemporaneously in free dialogue; speakers in the monologue condition recorded their narrative into a handheld recorder. The task for all speakers was simply to recall and relate the excerpts they had just seen.

The empirical literature on direct quotations suggests that most scholars tend to locate them in dialogue (e.g., Bangerter, Mayor, & Pekarek Doehler, 2011; Clark & Gerrig, 1990; Fox & Robles, 2010; Holsanova, 2006; Koike, 2001; Sams, 2007; Sidnell, 2006). However, to our knowledge, no one has compared dialogue with monologue or suggested an effect of visibility on direct quotations.

Our reviews of conversational facial gestures (e.g., Bavelas, Gerwing, & Healing, 2014) have consistently found only one experiment that used free dialogues: Chovil (1989, 1991) focused on addressees' facial mimicry (e.g., wincing) when listening to speakers tell about a close call they had in the past. In three dialogue conditions, the addressees listened to speakers' close calls either face to face, on the telephone, or through a partition. In the alone condition, the addressees listened to a recording of a close call about a horrific skiing accident (recorded by an earlier participant). Facial mimicry was significantly more likely in the three dialogue conditions than in the alone condition. There was also a significant effect of visibility; the face-to-face condition produced more facial mimicry than the group of conditions that precluded visibility (telephone, partition, and alone). However, because the latter group included the alone condition, the test did not control for dialogue versus monologue.

We predicted that, after controlling for visibility, both facial portrayals and direct quotations would occur at higher rates in the dialogue conditions than in the monologue condition. The results would also show whether Chovil's (1991) visibility effect on addressees' facial mimicry generalized to speakers' facial

portrayals when controlled for dialogue. In the case of direct quotations, there was no reason to predict any effect of visibility.

Method

Participants. Fifty-five female undergraduate psychology students signed up online in return for 1% toward their course grade. The posting described the purpose of the study as “to gain a further understanding of what people do when they communicate in different settings” and described their activities as talking about interesting and enjoyable things (e.g., movie scenes or games). They knew they would be videotaped and could control the research use of the video after they had seen it.

The order of conditions was randomized in advance, and the sign-up procedure prevented participants from signing up together and from knowing whether they would be participating in a dialogue or monologue. We excluded and replaced data from four participants: three were not sufficiently fluent in English, and one participant in the monologue condition rated herself as having felt she was talking to another person while doing the tasks (see Procedure). The final number of participants was therefore 51: 20 talked face to face (10 dyads), 20 talked on the telephone (10 dyads), and 11 individuals talked into a handheld recorder. Random assignment determined the roles of speaker and addressee in the dialogues.

Materials. The stimulus was a 2-minute, 45-s videotape containing two scenes from *Shrek 2*, an animated movie with a great deal of colloquial humor. The scenes included one humanoid figure (Shrek, an ogre) and two talking animals (Donkey and Puss in Boots, a cat). To minimize interaction with the experimenter (for the monologue condition), the speakers in all conditions had a set of cards with written instructions for each of their tasks.

Equipment. The experiment was conducted in the Psychology Department’s Human Interaction Laboratory suite, using three of its four Panasonic WV-CP474 color cameras and custom Panasonic special effects generators to capture a close-up of the speaker’s face, a circular insert of the addressee’s head and face in the face-to-face condition, and a narrow side view of both participants. (In [Figures 1](#) and [2](#), the transcript was superimposed on the side view.) During the experiment, the speakers viewed the movie scenes on a small color TV/VCR. In the telephone condition, speakers used a handheld telephone, with the telephone audio of both participants fed directly into the video recording system. Speakers in the monologue condition talked into a Sony TCM-900DV handheld mini-cassette recorder (8.67 × 3.57 × 11.28 cm) with an internal microphone; the analysis used the audio on the video recording. We digitized the

videos with Broadway ProDVD (www.b-way.com) and analyzed them on an 18-inch ViewSonic G90fb color monitor using either Broadway or ELAN (www.lat-mpi.eu/Tools/Elan; Brugman & Russel, 2004; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006).

Procedure. In the face-to-face condition, speakers and addressees interacted across a table in the main lab room. In the telephone condition, once the participants had signed their consent forms, the experimenter escorted the addressee to a nearby office while the speaker stayed in the main lab room. The speaker then phoned the addressee, and all their interactions were on the telephone. In the monologue condition, the speakers were alone in the main lab room, talking into the handheld recorder.

Before recording began, the experimenter gave an overview of the experiment, and the participants provided written consent, including consent to being videotaped for our later use. Then, to ensure the monologue condition would be as asocial as possible, the experimenter gave the speakers in all conditions four task-instruction cards and left the room. First, the pairs of participants got acquainted, and the individual in the monologue condition described herself into the recorder, then all participants did two unrelated pilot tasks.

Next, they read their respective task-instruction cards for the main task. All speakers began by going into an adjacent room to watch scenes from a movie twice “in order to remember as many details as possible.” All speakers received the same sheet of additional instructions that specified five parts of the movie they should include in their description (see Appendix A). Then they returned to the main room with instructions to describe the movie scenes as follows: In the face-to-face condition, each speaker was to “re-tell the scenes to your partner in detail.” In the telephone condition, each speaker was to phone and “re-tell the scenes to your partner in detail.” In the monologue condition, the speaker was to “describe the scenes into the tape recorder in detail.”

Afterward, the experimenters debriefed the participants, showed them their videotape, and each participant indicated in writing the allowable uses of their video (e.g., permission to view for analysis only, permission for a still photo in a journal article).

As part of the debriefing, the experimenter asked the monologue speakers to rate, on a written scale of 1 to 10, how much they felt as if they “were talking to another person” while doing the tasks (1=*not talking to anyone* and 10 = *always talking to someone*). This measure provided a manipulation check for whether this condition was effectively monologic in nature; $M = 3.55$, $SD = 2.33$. One participant was excluded from the analysis because her rating was between 8 and 9 (final $M = 3.05$, $SD = 1.59$). Although the ratings screened out the clearest case, all speakers knew they were being videotaped, so this

condition was still minimally social, corresponding to the seventh context in Table 1.

Analysis

Preparing a data set. First, we divided the video itself into 34 very short events (e.g., the cat begs for his life; the donkey objects to letting the cat join them), then identified the events that each speaker described; this created an inventory of all the data available for analysis. Because not every speaker described every scene in the movie, it was important to ensure comparable material across speakers as well as the maximum number of speakers in every condition. Our first criterion for selecting events from this inventory was the frequency with which speakers described them (i.e., we excluded events that only a few speakers described). The second criterion was that the number of speakers describing the event should be as equal as possible across the three conditions. These criteria resulted in 10 events (described in Appendix B) that the maximum number of speakers described and were most equally described across the three conditions.

*Facial portrayals.*³ This analysis used four of the above events that were most likely to include facial portrayals (i.e., the characters in the movie were reacting to something; see Appendix B). These criteria did not guarantee that all speakers described each of the four events, but the 31 speakers produced a total of 94 descriptions, which were the data analyzed here.

The operational definition of a *facial portrayal* was one that depicted a character in the movie (i.e., not the speaker's own facial reaction either to the movie or to the addressee; see examples in Figures 1 and 2). A facial portrayal could be any meaningful movement or combination of facial features that depicted one of the characters at a particular point in the movie. The portrayal could use specific facial muscles (e.g., smiling, wrinkling the nose in disgust, raising the eyebrows) and could also involve the speaker's face in other ways. That is, we included the use of the face as a prop (e.g., demonstrating very big eyes by a forming large circles with the hands in front of the eyes), any meaningful motion or orientation of the head (e.g., looking up or down as a character in the movie had done), or any meaningful motion or position of the eyes (e.g., rolling or widening the eyes).

Conversational facial gestures inherently involve motion; they are not discrete, singular events. This is particularly true for facial portrayals of characters in movies (vs. emotional expressions in a photo). Demonstrating a

³Full operational definitions for all measures in this article are available from the authors.



FIGURE 2 Example of a speaker's facial portrayal while describing a scene from *Shrek 2* in which Donkey objected to having the cat join him and Shrek on their trek. Frames 1 and 3 show the speaker's usual (i.e., baseline) smiling face, and Frame 2 shows her shift into a facial portrayal of Donkey's disapproving expression. The three frames cover 4.0 s. (The addressee's face has been blocked out in this figure.)

character at a particular point in the movie often included a rapid, complex, and overlapping sequence of facial actions. A typical example follows:

The speaker was demonstrating the cat's pleading silently with Shrek. While saying "He's got these big eyes, and he's like—," she briefly formed a semicircle with each hand in front of her eyes [the cat's big eyes], softened her face into a hopeful smile [the cat's expression], put her chin on her fists and tilted it forward and slightly up [looking up at Shrek], and flashed her eyes further up and to the side, showing more white [the cat's big eyes looking up at Shrek, who was much taller]. (1.99 s)

It would be difficult and questionable to divide these integrated and often overlapping actions into separate facial portrayals. Instead, the analysts treated this sequence as a single, unified portrayal of the how the cat looked and acted at that particular point in the scene. Rather than count the number of portrayals, our dependent measure was *the number of words that accompanied the portrayal*.

Two analysts applied the above operational definition to the 94 descriptions, which were divided into two groups. First, a subset had been identified as having no facial portrayal at all. A different, earlier set of analysts had screened each of the 94 descriptions for a possible facial portrayal of a character in the movie. In 51 descriptions, they saw no facial portrayal. The two analysts for the current analysis used the more precise definition above to confirm that there were no facial portrayals in any of them. Including these descriptions in any reliability calculation would have greatly inflated interanalyst agreement because both analysts would have agreed that zero words accompanied a facial portrayal—because there was no facial portrayal. (These descriptions were, of course, included in the results.) The two analysts focused on the 43 remaining descriptions. The analysis and reliability assessment proceeded event by event.

The two analysts worked together on a randomly selected 50% of the descriptions for each event and then independently on the other 50%.

First, they considered what a facial portrayal would be for a character in that event (e.g., the donkey's not wanting to take the cat with them). Then they watched the selection frame by frame without audio, noting if and when the speaker made a facial gesture that matched the expected portrayal. Analyzing first without audio avoided cross-modal distraction; the portrayals were often subtle, very quick facial motions, easily overshadowed or biased by what the speaker was saying. When the analysts noticed a potential facial portrayal, they checked the audio to ensure that it was the expected portrayal (e.g., not the speaker's own reaction). If so, then the analysts noted the exact accompanying words.

The calculation of reliability was word by word: Analysts had to agree whether each word the speaker used was or was not during the facial portrayal. They agreed on 404 of 445 words (90.78%) and resolved their disagreements. The kappa for agreement on words with a facial portrayal and words without a facial portrayal was .79 ("substantial" according to Landis & Koch's [1977] benchmarks). The dependent measure was a proportion: the number of words that were accompanied by a facial portrayal over the total number of words the speaker used to describe the four (or fewer) events.

Direct quotations. The data for this analysis were initially the 10 events in Appendix B. The operational definitions focused on the verbal characteristics of quotations as opposed to prosodic or visual cues. The two analysts worked from transcriptions created by a third person who had included only the speakers' words, without any commas, quotation marks, or other punctuation and who also grouped the transcripts by event with no indication of condition. Table 2 presents several examples of direct quotations from the data, contrasted with what analysts considered to be the use of descriptions (in Clark & Gerrig's [1990] sense) of the same event.

There were two stages of analysis. First, the analysts examined each event transcript for the presence of a *direct verbal quotation*, defined as any time the speaker said what one of the characters in the movie said or could have said in the situation. Common indications of direct quotations were introductory quotatives (e.g., "like," "he goes," "he asks"), using first person instead of third person (e.g., "I" instead of "he" for the character who was speaking), and using present tense (e.g., "he's like, 'Please don't hurt me'"). In the second stage, the analysts examined only the events with a direct quotation and underlined the exact words in the quotation.

The two analysts worked together on a randomly selected set of 2 of the 10 events for all speakers and then worked independently on a randomly selected set of four new events to assess reliability. Their reliability for the first stage (locating events with a direct quotation) was 121 agreements out of 125 decisions (96.8%). After resolving their four disagreements, they did the

TABLE 2
Examples of Direct Quotations Contrasted With Descriptions for Events Being Recalled

<i>Event as It Occurred in the Movie</i>	<i>Participants' Direct Quotations</i>	<i>Participants' Descriptions^a (e.g., Indirect Quotations)</i>
As Shrek and the Donkey are about to leave on their trek, the cat says "Stop Ogre, I have misjudged you. On my honor, I am obliged to accompany you until I have saved your life as you have spared me mine."	"Puss in Boots is standing there and he is like, 'Please wait Ogre, I have misjudged you.'" "The cat asks them to stop, and he's like, 'I have misjudged you,' um, 'You spared me my life, now I will help you save yours' or something."	"The cat asks to go with Shrek because he spared him his life, so he wants to help save his." "And then the cat says that it would be in his honor to go with them because he has to repay Shrek for saving his life."
The donkey doesn't want the cat to come along and says "I'm sorry, the position of annoying talking animal has already been taken. Let's go, Shrek."	"And Donkey's like, 'No, we already have one annoying talking animal – that post's taken.'" "And then Shrek, Donk- the Donkey doesn't want the cat to come along and says something about like, 'The talking animal position in this group is already taken, thank you very much,' or something like that."	"Oh, but Donkey really doesn't want the cat to go with them so he tries to persuade Shrek to let him just stay there." "And, um, the Donkey doesn't want Shrek to take the cat because, I don't know, he just doesn't want him along."

^aThese are descriptions in Clark and Gerrig's (1990) sense, that is, verbal symbols with their conventional meanings.

second stage separately for the same four events, with each analyst locating the exact words that were in the direct quotations within those event descriptions. They agreed on 386 of 391 words (99.2%). Their resolved decisions completed the analysis of 6 of the 10 events for all speakers. Because their independent reliability was so high for both stages, one analyst completed the last four events on his own. After the analysis, four events were excluded because they yielded only four direct quotations (all in dialogue conditions). The final dependent measure was the number of events that included a direct quotation as a proportion of the number of events (up to a maximum of six) that the speaker described. Cohen's kappa for whether an event had a direct quotation or not was .93.

Results

Facial portrayals. The dependent measure was the proportion of words that were accompanied by a facial portrayal, which controlled for how much a speaker talked about each event. As shown in Table 3, speakers in the face-to-face dialogue condition used facial portrayals with almost a third of their words ($M = .30$, $SD = .15$), speakers in the telephone dialogue condition used them half of that amount ($M = .14$, $SD = .14$), and speakers in the monologue condition rarely used them ($M = .01$, $SD = .03$). The ANOVA indicated differences among these mean proportions, and the post-hoc tests showed no significant difference between the face-to-face and telephone conditions. The mean of the face-to-face condition was significantly higher than the mean in the monologue condition, but the mean of the telephone condition was not quite significantly higher than the mean of the monologue condition ($p = .052$).

We also report the results of two simple linear regression tests and two multiple linear regression tests in Table 3:

- *Dialogue alone:* How much variance did dialogue account for as the only predictor?
- *Visibility alone:* How much variance did visibility account for as the only predictor?
- *Dialogue controlled for visibility:* After partitioning out the variance accounted for by visibility alone, how much of the remaining variance did dialogue account for?
- *Visibility controlled for dialogue:* After partitioning out the variance accounted for by dialogue alone, how much of the remaining variance did visibility account for?

As predicted, the regressions indicated a significant effect of dialogue on facial portrayals after controlling for mutual visibility (i.e., face-to-face and

TABLE 3
Results of Experiment 1: Effects of Dialogue and Visibility on Facial Portrayals and Quotations

Variable	Descriptive Statistics				One-Way ANOVA		Regressions			
	Face to Face		Tape Recorder		Main Effect, MSE, R^2	Pairwise Comparison ^{a,b}	Effect of Dialogue		Effect of Visibility	
	M (SD)	n	M (SD)	n			Alone	Controlled for Visibility	Alone	Controlled for Dialogue
Proportion of facial portrayals	.30 (.15)	10	.14 (.14)	11	$F_{(2,28)} = 14.90^{***}$ MSE = .01 $R^2 = 52\%$	FF > TR	$F_{(1,29)} = 16.62^{***}$	$F_{(1,28)} = 5.90^*$	$F_{(1,29)} = 20.45^{***}$	$F_{(1,28)} = 8.74^{**}$
Proportion of direct quotations	.82 (.23)	10	.55 (.31)	16	$F_{(2,38)} = 27.58^{***}$ MSE = .06 $R^2 = 66\%$	FF > TR PH > TR	$F_{(1,29)} = 41.13^{***}$	$F_{(1,28)} = 22.15^{***}$	$F_{(1,29)} = 19.09^{***}$	$F_{(1,28)} = 6.39^*$
							$R^2 = 36\%$		$R^2 = 41\%$	$R^2 = 40\%$

^a Because of nonhomogenous variance, pairwise comparisons were tested with Dunnett's T3.

^b FF = face-to-face dialogue condition; PH = telephone dialogue condition; TR = tape recorder monologue condition.

* $p < .05$, ** $p < .01$, *** $p < .001$. All tests are two-tailed.

telephone dialogues versus monologues). There was also a significant effect of visibility after controlling for dialogue (i.e., face-to-face versus telephone and monologue), so both dialogue and visibility accounted for a unique amount of the variance in facial portrayals.

Direct quotations. The dependent measure was the proportion of event descriptions with a direct quotation, which controlled for how many events the speaker described. As shown in Table 3, speakers in the face-to-face condition were very likely to include a quotation in their event descriptions ($M = .82$, $SD = .23$). Speakers in the telephone condition used fewer quotations ($M = .55$, $SD = .31$), and those in the monologue condition rarely used quotations ($M = .06$, $SD = .16$). The ANOVA indicated a difference among means, and post-hoc tests showed that (1) speakers in both the face-to-face and telephone dialogue conditions used quotations at a significantly higher rate than those in the monologue condition and (2) the face-to-face and telephone conditions were not significantly different from each other. Regardless of the number of events each speaker chose to retell, the proportion that included a quotation (versus pure description in Clark & Gerrig's [1990] sense) was highest in the two dialogue conditions. The two multiple linear regressions reported in Table 3 showed a significant effect of dialogue on direct quotations after controlling for mutual visibility, as well as a significant effect of visibility after controlling for dialogue.

Discussion

As predicted, dialogue increased the proportions of both kinds of demonstrations. Even after controlling for visibility, dialogue accounted for a significant amount of variance in speakers' direct quotations and facial portrayals. The results for facial demonstrations are consistent with Chovil's (1991) findings, which extends the effect of dialogue on demonstrations to include speakers' facial portrayals as well as addressees' facial mimicry. The results for direct quotations are entirely new to the literature and show that Clark and Gerrig's (1990) prototypic demonstration is sensitive to dialogue versus monologue.

Although visibility was a control variable in this experimental design, there were two noteworthy effects. First, visibility led to a significantly higher rate of facial portrayals, which replicates Chovil's (1991) finding for addressee's facial mimicry, even after controlling for dialogue. Second, the effect of visibility on direct quotations (again, after controlling for dialogue) is surprising; we return to the puzzle of visibility in the General Discussion.

It is important to point out that the facial mimicry in Chovil (1991) and the facial portrayals studied here are only two kinds of facial gestures. Descriptive studies of conversational facial gestures have shown that faces serve numerous other functions in dialogue. These functions are as diverse as eyebrow actions

that stress a word or phrase (Chovil, 1991/1992; Ekman, 1979) and smiles that act as back-channels (Brunner, 1979). Therefore, although facial mimicry and facial portrayals are clearly demonstrations that increase in dialogue, only wider research would support generalizations about other conversational facial gestures.

We also note that facial portrayals and direct quotations had identical patterns in the regression analyses (i.e., each showed both a dialogue and a visibility effect). This might seem to imply that these two demonstrations are themselves linked in some way, either occurring simultaneously or one causing the other. However, there are several reasons to be cautious about such a hypothesis. First, these results are aggregate statistical effects; they provide no information about specific, momentary local co-occurrences. Second, the data sets were overlapping but different. In the movie events chosen for the analysis of direct quotations, one of the characters had to say something; in the events analyzed for facial portrayals, one of the characters had to be reacting to something (and did not necessarily say anything). For example, one event consisted of the cat character pleading silently with Shrek. This event elicited a high rate of facial demonstrations but only two quotations. Based on the present data, a more parsimonious explanation for the parallel statistical effects on facial portrayals and direct quotations is found in the nature of the material the speakers were relating. These particular movie excerpts were chosen because they included distinctive faces and clever dialogue. That is, these excerpts were good stimuli for eliciting depictions of the characters with a facial portrayal, a direct quotation, or both, depending on what the movie character was doing in each particular event. Note that pure description (in Clark & Gerrig's [1990] sense) was always still an option, as illustrated by speakers in the monologue condition.

EXPERIMENT 2: HAND GESTURES AND FIGURATIVE REFERENCES

The purpose of this study was to replicate two unexpected effects of dialogue and visibility on the rates of gesturing and figurative references found in Bavelas et al. (2008), using previously unanalyzed data from that experiment. First, the effect of dialogue on figurative references in Bavelas et al. was serendipitous and new to the literature. Most experiments on figurative language involve participants reading and rating written examples (e.g., Colston & Katz, 2005). Boerger (2005) investigated figurative language in dialogues that varied in visibility (i.e., comparing face-to-face, partition, intercom, and email conditions). However, addressees who were face to face could also see the objects the speakers were describing, whereas those in the other conditions could not, which makes these results hard to interpret. Garrod and Anderson (1987) and Garrod and Doherty (1994) did not vary either dialogue/monologue or visibility, but their dyads

working together on a maze task sometimes spontaneously aligned on the use of “figural descriptions,” which are similar to the figurative references found in Bavelas et al. Finally, it might be relevant that the stimulus in our previous experiment was difficult to describe verbally, which could have favored the use of figurative language as an option. For all the above reasons, we sought to replicate the effect of dialogue on figurative language with a different stimulus. As shown in Figure 3, the stimulus was a drawing of several geometric shapes connected by a line; all these shapes had familiar conventional terms, so figurative language was neither necessary nor a better option. Still, we predicted that the rate of figurative language would be higher in the dialogue conditions than in the monologue condition.

Second, although the purpose of Bavelas et al. (2008) was to separate dialogue from visibility, the incidental finding of no difference in overall rate of hand

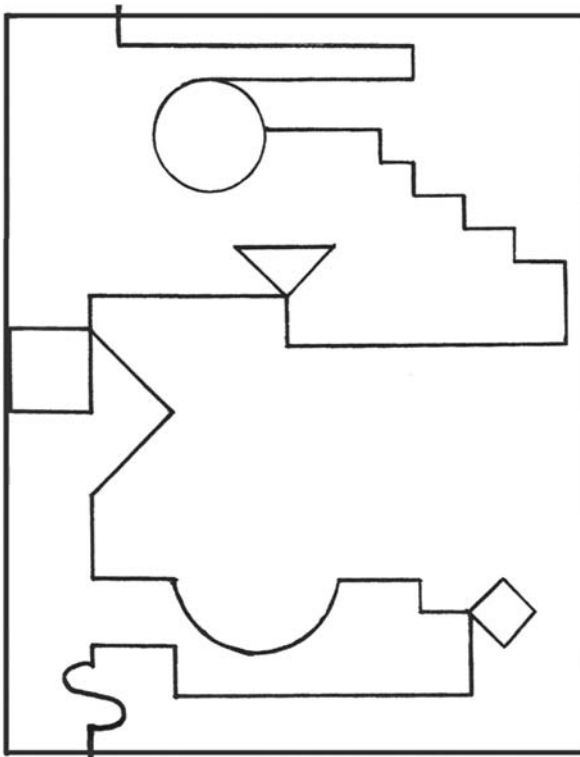


FIGURE 3 The stimulus for Experiment 2, the abstract line drawing.

gestures between the visible and not visible conditions was an anomaly in the literature up to that date (see Bavelas et al., 2008, Table 1, p. 497). With two exceptions (Bavelas et al., 1992; Rimé, 1982), all earlier experiments had found a significant visibility effect. Bavelas et al. (2008) added a third exception and pointed out (pp. 511–512) that all three exceptions had an addressee who was a real participant, interacting freely, whereas in all the experiments that had found a visibility effect, the addressee was interacting within constraints (e.g., was a confederate or the experimenter). Since 2008, the number of visibility experiments has almost doubled—and has confirmed this pattern: Six experiments that used free dialogues (i.e., the first context in Table 1) have found no significant difference between the visible and not-visible conditions (Bavelas et al., 1992, 2008; de Ruiter et al., 2012; Holler et al., 2011; Pine et al., 2010; Rimé, 1982). Seven experiments with constrained dialogues (i.e., the third context in Table 1) have found a significant difference (Alibali et al., 2001; Cohen, 1977; Cohen & Harrison, 1972; Emmorey & Casey, 2001; Krauss et al., 1995; Mol et al., 2009a,b). For a detailed examination of this discrepancy, see Bavelas and Healing (2013). We therefore predicted that the present data would replicate the growing body of evidence with free dialogues and would find no visibility effect: The overall rates of gestures in the face-to-face and telephone dialogues would be higher than in the monologues and the two dialogues would not be different from each other (i.e., a dialogue effect and no visibility effect). It is important to emphasize that the predicted lack of a visibility effect refers only to *overall gesture rate*. Recent research shows a wide variety of ways in which visibility between participants significantly influences more specific features of gestures (see review in Bavelas & Healing, 2013).

In the dialogue conditions (as in Bavelas et al., 2008), we took steps to ensure a free dialogue by giving the addressee a reason to be more than a passive listener while the speaker simply described a picture that the addressee could not see: Both participants knew that the addressee would later have to identify this particular picture from an array of four similar pictures. Therefore, as is common in referential communication tasks (Glucksberg & Krauss, 1967; Yule, 1997), the speakers in the dialogue conditions were engaged in *social speech*, taking the addressee's knowledge into account, whereas the speakers in the monologue condition were engaged in *nonsocial speech*, speaking from their own perspective and focused on the quality of their individual performance. Therefore, in the monologue condition, there was no mention of anyone who might hear the recording later, because even imaginary audiences can elicit gestures (Bavelas, Kenwood, Johnson, & Phillips, 2002). Notice that the addition of a simple test for the addressee in the two dialogue conditions could not plausibly increase the rate of figurative references (because more precise conventional terms were readily available) or of gestures (because the addressee would not see them in the telephone dialogues).

Method

Participants. Initially, 61 first-year psychology students signed up online to participate in return for .5% toward their course grade. The order of experimental conditions was randomly assigned in advance, but the online sign-up procedure prevented participants from knowing which condition they would be in or from signing up with a friend. During data collection, we replaced one dyad because the experimenter made a procedural error. We also replaced the following: four individuals in the monologue condition who reported they had imagined talking to someone (e.g., the experimenters) and whose language included interactive phrases such as “you know” or “sorry,” two dyads who explicitly reported trying not to gesture, and one individual who reported both imagining an addressee and trying not to gesture. The final number was therefore the planned 50 participants: 20 in the face-to-face condition (forming 10 dyads), 20 in the telephone condition (10 dyads), and 10 individuals in the monologue condition. The roles of speaker and addressee were randomly assigned.

Materials. The stimulus was the black and white line drawing shown in [Figure 3](#), laminated onto an approximately $8\frac{1}{2} \times 11$ -inch cardboard sheet. A small stand held the picture up facing the speaker. The addressees later saw a large placard with four digitally edited versions of the same line drawing, one of which was identical to the original stimulus.

Equipment. The experiment took place in the Human Interaction Laboratory suite. We used three Panasonic WD-D5000 color cameras and a custom Panasonic special effects generator system to capture a full view of the speaker from the front and from the side plus a circular insert of the addressee’s head and face in the face-to-face condition. The speaker used a handheld telephone with the audio from both participants fed directly into the video system. In the monologue condition, speakers spoke into a tape recorder on the table in front of them using a handheld microphone. All analyses used the audio on the video recording. Broadway ProDVD (www.b-way.com) was used to digitize the analogue video into AVI format and also for frame-by-frame analysis and precise repetition of selected sections. Viewing for analysis was on an 18-inch ViewSonic GS790 color monitor.

Procedure. The seating arrangements for the three conditions were the same as in Experiment 1. Before recording began, the participants gave consent in writing to being in the experiment and being videotaped. The experimenter then gave an overview of instructions for the main tasks and provided a written copy of instructions as well. After getting acquainted (in the dialogue conditions) or telling a bit about themselves (in the monologue condition), the speakers

described the line drawing as well as the other picture analyzed in Bavelas et al. (2008), in counterbalanced order. In all three conditions the speakers were to take the picture out of the folder, place it in the vertical stand directly in front of them, and then describe it in “the clearest and most detailed way that you can.” In the two dialogue conditions, the experimenter added that they could “talk and ask questions whenever you need to” and that the addressee would later choose the described picture out of four similar pictures. In the monologue condition, there was no mention of any addressee. When they announced they were done, the experimenter re-entered and presented the addressee with the four options to choose from. The addressee was told that the choice was correct (which was true in all but one case). The speaker did not see the four options. They then repeated this process for the second picture.

During debriefing, the experimenter asked each speaker in the monologue condition whether he or she had imagined an audience while describing the picture (i.e., the sixth context in Table 1). As described in Participants above, data from five speakers were replaced because two lines of evidence indicated they had been imagining an audience. Although our criteria screened out the clearest cases, all speakers knew they were being videotaped. Therefore this condition was still minimally social (i.e., the seventh context of our proposed continuum).

At the conclusion of the experiment, all participants received an explanation of the study, had the opportunity to ask questions, viewed the videotape of their participation, and indicated in writing which uses of the videos they consented to (e.g., permission to view for analysis only, permission for viewing by professional audiences).

Analysis

Hand gestures. The data were the first minute of each speaker’s description. The analysts made two decisions, locating hand gestures and then dividing them into separate gestures when appropriate. Each analyst examined all the speaker’s hand and arm movements to locate *conversational hand gestures*, that is, purposive and meaningful hand movements that were synchronized with speech and that illustrated the speaker’s words (i.e., not self- or object-oriented adaptors; Ekman & Friesen, 1969). See the example in Figure 4. When there was an apparently continuous period of gesturing, the analyst decided whether or not to separate it into discrete gestures, based on either timing (more than two frames or 66 ms between the gestures) or meaning (e.g., gesturing two different features in the drawing). See the example in Figure 5.

As in Bavelas et al. (2008), the analysis ensured that having both hands free in the face-to-face condition could not artificially increase the number of gestures. The analysts counted a gesture with both hands as a single gesture unless the



FIGURE 4 Example of a conversational hand gesture, which began in frame 1 and ended in frame 2. While demonstrating one of the diagonal lines on the left side of the drawing in Figure 3, the speaker made a purposive downward diagonal movement (shown by the arrow) that was synchronized in timing and meaning with her words, “going downwards.” (Total time = 2.4 s.)

referents were distinctly different (which was rare). For example, the two-handed circle in frame 3 of Figure 5 was one gesture.

One person analyzed all the data, and then a second analyst, working independently, applied the same operational definitions and procedures to a stratified random sample of 20% of the data (two speakers from each condition). For locating periods of gesturing, they agreed on 98 of 104 their decisions (94.2%). That is, six times one analyst located a gesture that the other analyst did not, and they resolved these six decisions. For separating gestures, they agreed on 90 of 97 decisions (92.8%) and again resolved their disagreements. These



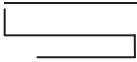


FIGURE 5 An example of two discrete hand gestures while the speaker was demonstrating a horizontal line going into a circle (near the top of Figure 3). Frame 1 shows her hand pulled back in the preparatory phase. The arrow in frame 2 shows the horizontal line she drew. In frame 3, she made a circle at the end of the line and held it in place. The speaker made one smooth, continuous motion without retracting her left hand to a resting position between the two gestures. However, both her hands and words clearly referred to two different features, so the analysts treated these as two gestures. These gestures took 3.6 s. (The addressee’s face has been excluded in this figure.)

decisions are not amenable to kappa, but the correlation between the two analysts for the number of gestures they found for each speaker was $r = .98$. The final data consisted of their consensus decisions plus the first analyst's decisions regarding the remaining 24 speakers. The two measures of overall gesture rate were the total number of gestures in the first minute and the rate per 100 words.

Figurative references. Our previous experience suggested that the rate of figurative reference would be lower than the rate of gesturing (Bavelas et al., 2008). Therefore, rather than using only the first minute, we analyzed all the references that speakers made to the nine distinct geometric features throughout their entire description. There were two stages of analysis. The first stage was to locate each noun or noun phrase that the speaker used to refer to one of the nine features. The second stage was to decide whether each reference was literal or figurative. The operational definitions were based on Crystal's (2001) definition of figurative language, "an expressive use of language where words are used in a non-literal way to suggest illuminating comparisons and resemblances" (p. 116). Thus, a *literal reference* was the geometric name or a denotative description of the feature, and a *figurative reference* was verbal imagery, including metaphor, simile, or analogy. Table 4 presents three features from the drawing with examples of both figurative and literal references taken from the data.

One analyst analyzed all the data, first locating the references and then deciding whether each was literal or figurative. A second analyst used the same operational definitions and procedures for both stages and worked independently on a stratified

TABLE 4
Examples of Figurative and Literal References to Three Features of the Line Drawing

Features	Figurative References	Literal References
	"old-school cannon" "cylinder" "long strip" "zed thing"	"goes down and it goes right and then it goes down" "down right down" "goes perpendicularly to the right" "straight parallel lines"
	"martini glass without the bottom part" "wine glass shape" "cup and the stem bit"	"triangle" "upside down triangle" "triangle and a line"
	"little circle dip" "ditch" "these things where the skateboarders skate" "halfpipe"	"half circle" "concave half circle" "semicircle" "curve"

random sample of 20% of the data (two from each condition). In the first stage they agreed on 152 of 162 references to the nine features (93.8%) and resolved the 10 disagreements. In the second stage they agreed on 153 of 156 decisions (98.1%) and resolved the three disagreements. Cohen's kappa for the second decision (literal vs. figurative) was .96. Their consensus agreements plus the first analyst's decisions on the remaining data yielded the raw frequencies that went into the final measure, which was the proportion of a speaker's total references that were figurative. (This measure is more precise than the rate of figurative references per 100 words used in Bavelas et al. [2008], which included the words that speakers used to talk about other aspects of the drawing or the task.)

Results

Hand gestures. Table 5 shows all results for the rates of gestures per minute and per 100 words. One speaker in the monologue condition was identified by SYSTAT13 (Systat Software Inc., San Jose, CA, USA) as an outlier (studentized residual = 3.204) and was therefore removed from statistical analysis. Speakers in the face-to-face dialogue condition used gestures at the highest rate per minute ($M = 21.60$, $SD = 7.78$), followed closely by speakers in the telephone dialogue condition ($M = 19.00$, $SD = 10.19$), and then by speakers in the monologue condition at a much lower rate ($M = 7.33$, $SD = 7.02$). The ANOVA indicated a significant difference among these three means, and post-hoc tests confirmed (1) that the means of the face-to-face and telephone dialogue conditions were each significantly higher than the mean of the monologue condition and (2) that the means of the face-to-face and telephone dialogue conditions were not significantly different from each other. The results for gesture rate per 100 words were the same (see Table 5). Thus, even when the speaking rate was held constant, the proportion of gestures was still significantly higher in the two dialogue conditions than in the monologue condition. Finally, the two multiple linear regressions reported in Table 5 showed that dialogue accounted for a significant amount of variance in overall gesture rate, even after controlling for visibility, but visibility did not account for a significant amount of variance after controlling for dialogue.⁴

Figurative references. Table 5 also includes the results for the proportion of figurative references, which controls for the total number of references to features of the drawing. One speaker in the monologue condition was identified

⁴Bavelas et al. (2008) reported a regression for the overall rate of hand gestures that showed a significant visibility effect. However, this regression did not control for dialogue. A subsequent reanalysis added a control for dialogue, which eliminated the apparent visibility effect. The reanalysis is congruent with the post-hoc finding of no significant difference between the face-to-face and telephone dialogues (2008, p. 512) and with the present results.

TABLE 5
Results of Experiment 2: Effects of Dialogue and Visibility on Hand Gestures and Figurative Language

Variable	Descriptive Statistics <i>M (SD)</i>			One-Way ANOVA		Regressions			
	Face to Face <i>n</i> = 10	Telephone <i>n</i> = 10	Tape Recorder <i>n</i> = 9	Main Effect, <i>MSE, R²</i>	Pairwise Comparison ^{a,b}	Effect of Dialogue Alone	Controlled for Visibility	Effect of Visibility Alone	Controlled for Dialogue
Gestures per minute	21.60 (7.78)	19.00 (10.19)	7.33 (7.02)	$F_{(2,26)} = 7.48^{**}$ <i>MSE</i> = 72.02 $R^2 = 37\%$	FF > TR ± 9.69 PH > TR ± 9.69	$F_{(1,27)} = 14.78^{***}$ $R^2 = 35\%$	$F_{(1,26)} = 8.95^{**}$	$F_{(1,27)} = 4.64^*$ $R^2 = 15\%$	$F_{(1,26)} = .47$
Gestures per 100 words	14.86 (4.44)	12.98 (6.05)	6.28 (5.61)	$F_{(2,26)} = 6.52^{**}$ <i>MSE</i> = 29.16 $R^2 = 33\%$	FF > TR ± 12.33 PH > TR ± 12.33	$F_{(1,27)} = 12.62^{***}$ $R^2 = 32\%$	$F_{(1,26)} = 7.31^*$	$F_{(1,27)} = 4.65^*$ $R^2 = 15\%$	$F_{(1,26)} = .60$
Proportion of figurative references	.38 (.05)	.41 (.10)	.22 (.10)	$F_{(2,26)} = 13.61^{***}$ <i>MSE</i> = .01 $R^2 = 50\%$	FF > TR ± .10 PH > TR ± .10	$F_{(1,27)} = 25.95^{***}$ $R^2 = 49\%$	$F_{(1,26)} = 23.28^{***}$	$F_{(1,27)} = 1.66$ $R^2 = 6\%$	$F_{(1,26)} = .68$

^a Pairwise comparisons are reported as 95% confidence intervals (HSD) around the differences between means.

^b FF = face-to-face dialogue condition; PH = telephone dialogue condition; TR = tape recorder monologue condition.

* $p < .05$, ** $p < .01$, *** $p < .001$. All tests are two-tailed.

by SYSTAT as an outlier (studentized residual = 3.077), so this speaker's score was removed from statistical analysis. As shown in Table 5, speakers in the face-to-face and telephone dialogues used similar proportions of figurative references ($M = .38$, $SD = .05$ and $M = .41$, $SD = .10$, respectively). Speakers in the monologue condition used the lowest proportion of figurative references ($M = .22$, $SD = .10$). The ANOVA indicated a difference among these means, and post-hoc tests confirmed that the mean proportions in the telephone and face-to-face conditions were (1) significantly higher than in the monologue condition and (2) not significantly different from each other. Note that holding constant the number of references that each speaker chose to make permits a comparison between the proportion that were literal (i.e., description in Clark & Gerrig's [1990] sense) versus figurative. The results confirm that this proportion changed as a function of experimental condition. Finally, the two multiple linear regressions reported in Table 5 showed that, as with gestures, dialogue accounted for a significant amount of variance in the proportion of figurative references even after controlling for visibility, but visibility did not account for a significant amount of variance after controlling for dialogue. Dialogue alone accounted for 49% of the variance.

Discussion

The results replicated the Bavelas et al. (2008) findings that dialogue itself—whether in person or on the phone—produced a higher rate of gesturing and a higher proportion of figurative references than did the monologue condition, even after controlling for visibility. Although these data were from the same speakers as the 2008 results, which limits their generalizability, the two pictures differed in an important way. The 2008 picture of an 18th century dress was one for which speakers had virtually no technical vocabulary. In contrast, the abstract line drawing consisted of familiar geometric features for which speakers would have a readily available conventional vocabulary, so there should be much less need to use either gestures or figurative references. In spite of this difference, the effects of dialogue and mutual visibility were the same for both pictures.

The results also show that using a telephone did not affect the overall gesture rate, for example, lowering it because of having only one hand free or increasing it by making more uni-manual gestures to convey the same information as a bi-manual gesture would. Either of these possibilities would have affected the telephone and microphone conditions equally. Yet, as in the 2008 experiment, these two conditions were significantly different, and the telephone condition was not significantly different from the face-to-face condition. The regressions also showed that visibility had no effect on overall gesture rates after controlling for dialogue. The absence of a significant difference between the means of the two

dialogue conditions replicates the findings of the six other experiments (cited above) that also used free dialogues. We should re-emphasize that this finding refers only to the overall rate of gesturing. At least four of the previous six experiments that found no effect of visibility on their omnibus rate measure also reported significant effects of visibility on more specific features of gestures (Bavelas et al., 1992, 2008; de Ruiter et al., 2012; Holler et al., 2011).

It is noteworthy that, as in Bavelas et al. (2008), gestures and figurative references showed identical effects. As Corts (2006, p. 212) pointed out, gestures and figurative references serve similar functions:

Although treated separately in the literature, figurative reference and gestures often serve similar functions; they provide images, they add to the message above and beyond what is explicitly presented in speech, and they describe complex ideas in simpler terms than are readily available in literal language.

That is, both gestures and figurative references are demonstrations that are particularly well suited to providing the images required to convey the pattern of a drawing to an addressee, just as direct quotations and facial portrayals were well suited to the characters in *Shrek 2*.

GENERAL DISCUSSION

We set out to expand knowledge of demonstrations as a distinct class of communicative acts. We first pointed out that direct quotations, facial portrayals, hand gestures, and figurative language all fit the identifying criteria for demonstrations; that is, each relies on a resemblance to its referent rather than an arbitrary or conventional relationship, and each presents a selective depiction of its referent.

The main hypothesis was that they also shared a sensitivity to dialogue over monologue. These two experiments comparing face-to-face dialogues, telephone dialogues, and monologues confirmed that the rate of demonstrations was significantly higher in dialogues than in monologues and that this effect was independent of any effect of visibility. Experiment 1 extended the effect of dialogue on demonstrations to two new demonstrations. The results for facial portrayals replicated Chovil's (1991) findings with a second kind of facial gesture. The results for direct quotations are new to the literature and are important, in part, because direct quotations were the first systematically identified demonstration (Clark & Gerrig, 1990). Experiment 2 replicated the original effect of dialogue on hand gestures and figurative references (Bavelas et al., 2008) using a stimulus with different characteristics. In both experiments the effect sizes for dialogue were substantial. The four demonstrations studied

here differed in many respects, notably that hand gestures and facial portrayals are visible (nonverbal), whereas figurative references and direct quotations are audible (verbal), yet all four occurred at higher rates in dialogue than in monologue. Their diversity makes an even stronger case for the claim that free dialogue increases demonstrations.

Note that the dependent measures precluded the possibility that the monologue condition exerted a general dampening effect that suppressed descriptions (in Clark & Gerrig's [1990] sense) as well as demonstrations. Specifically, the denominator for each measure was the number of words or events that the speaker chose to use: the proportion of words accompanied by a facial portrayal, the proportion of events that included a direct quotation, the proportion of references that were figurative, and the number of gestures per 100 words. Thus, even after holding constant the amount that the speaker talked, the rate of demonstrations in the monologue condition was significantly lower.

Limitations

We only studied one exemplar of each type of demonstration, which limits the generalizability of our findings. That is, our analysis focused on the direct quotations of fictitious characters in a movie (e.g., not quotations from real interactions). Similarly, figurative language occurs in many other settings and on different topics than describing a picture. We focused only on how speakers used their face to portray fictitious characters, whereas descriptive studies of facial gestures have shown that faces serve a wide variety of other semantic and syntactic functions in conversation. Finally, as emphasized earlier, we studied the overall rate of all gestures and not more specific features of these gestures.

It is important to note that we do not conclude from these results that dialogue elicits demonstrations indiscriminately. In both experiments, the stimulus given to the speakers was chosen for the likelihood that it would elicit the demonstrations of interest in that particular experiment (e.g., the abstract line drawing would have been much less likely than the excerpts from *Shrek 2* to elicit quotations and facial displays). For the same reason, we would caution against using the present data to infer a temporal or causal relationship either between quotations and facial portrayals or between hand gestures and figurative references. In both experiments, it is more likely that the stimulus was the direct cause of the kinds of demonstrations the speakers used.

However, we would definitely not argue against multimodality in general. In Experiment 1, the speakers often used a highly integrated set of demonstrations to “become” the character they were portraying; for example, they combined facial and hand gestures in [Figure 1](#) (frame 2) and a quotation with

a facial portrayal in [Figure 1](#) (frame 3) and [Figure 2](#) (frame 2). Recall that speakers also occasionally used their hands as part of a facial portrayal of the cat's large eyes. Studies of such re-enactments reveal that multimodality is more than just words and hand gestures (e.g., Holsanova, Johansson, & Holmqvist, 2008), and the co-occurrence, timing, and other relationships among these complex demonstrations are worthy of further research. Setting up such studies requires addressing two technical issues. First, as noted above, the stimulus must facilitate all the demonstrations of interest. This can be difficult; for example, the *Shrek 2* excerpts elicited quotations, facial portrayals, and hand gestures; however, they elicited few figurative references beyond "puppy-dog eyes." Second, to have analyzable data, the split-screen recording must include both a close-up of the speaker's face (e.g., [Figures 1 and 2](#)) and a wider front view of the same person to capture gestures (e.g., [Figures 4 and 5](#)). Capturing both would require two cameras with different foci, both facing the speaker. Most studies have to choose whether they want faces or hands; in Experiment 1 we chose faces, and in Experiment 2 we chose hands.

Another possible limitation concerns the possibility of analysts' bias. The analysts for both experiments were all members of the research team and were therefore to some degree aware of the purpose of the experiments. Moreover, the experimental condition was obvious on each video. Still, there are reasons to doubt the extent to which bias could have contributed to the results. First, bias was unlikely for direct quotations and hand gestures: The transcripts used for the quotations analysis did not reveal the experimental condition. The possibility of bias on hand gestures was tested in Bavelas et al. (2008, p. 507), where completely naive analysts who each saw only one experimental condition produced the same results as analysts who were aware of both condition and hypothesis.

A second point applies to all the analyses: We propose that microanalysis with high interanalyst reliability tends to preclude the possibility of bias. The decision each analyst makes refers to a specific behavior at a specific point in time. Therefore, it is necessary to agree, for example, on exactly when each gesture occurred, not simply that a certain number of gestures occurred. Analysts could only achieve reliability at this micro-level, where we assessed it, by basing each decision on their common analysis rules rather than on their hypothesis. Furthermore, bias in a microanalysis would mean that each analyst would have unconsciously tended to make some of these decisions in a way that nudged the data toward a significant effect of dialogue that was independent of visibility (i.e., a complex pattern of means and standard deviations over the three conditions). In short, bias in support of a complex hypothesis is antithetical to high reliability when making the numerous precise decisions required in a microanalysis.

Effects of Mutual Visibility

We varied mutual visibility in these studies solely as a control variable, to separate the variables of social interaction and visibility, that is, to show that the effect of dialogue on hand and facial gestures in the face-to-face condition was not simply because the addressee could see them. As predicted, the post-hoc comparisons of the face-to-face and telephone conditions were not significant for any of the four demonstrations. Indeed, the only significant visibility effects (after controlling for dialogue) were in the regression analyses of direct quotations and facial portrayals in Experiment 1.

Altogether, these results produced new questions about visibility itself because they are clearly not consistent with a common-sense distinction between visible (nonverbal) and audible (verbal) demonstrations. That is, visibility increased the rate of facial portrayals but not of gestures, and it increased the proportion of direct quotations but not of figurative references. Moreover, even for hand gestures, which are the most-studied visible demonstration, the effect of mutual visibility in free dialogues is different for the overall rate (where there is no visibility effect) than it is for the rates of specific features (where there are definitely visibility effects). See Bavelas and Healing (2013, pp. 76–78 and Table 5) for a summary of the experiments showing that speakers modify specific features of their gestures for the benefit of addressees who will see them.

Possible Explanations or Mechanisms

Our first theoretical goal has been to treat demonstrations as a class and to provide evidence for a direct effect of dialogue on four members of this class. The present results, combined with those in Bavelas et al. (2008), justify beginning to consider more detailed explanations or mechanisms. That is, what characteristics of dialogue are responsible for this effect? Below are five plausible possibilities; they are not mutually exclusive, and all require further development and research.

1. The simplest approach would be to define dialogue and monologue as fundamentally different kinds of language use. Monologue would be a reduced and circumscribed form that relies mainly on description, whereas dialogue would combine Clark and Gerrig's (1990) description, demonstration, and indication. This possibility is consistent with our interpretation of McNeill's (2005, p. 82) later growth-point theory of hand gestures, in which the "social interactive context" (presumably including dialogue) has its influence at the earliest possible stage of production. However, a monologue–dialogue dichotomy would have to handle a continuum such as proposed in Table 1, that is, would have to specify where dialogue stops and monologue begins.

2. Linell (2005, Chapter 2) proposed several functional distinctions between written (monologic) and spoken (dialogic) language. He proposed that written language can stand alone; readers can understand it in nonsocial contexts because it is relatively explicit, autonomous, and context-free. In contrast, spoken language (by which Linell meant face-to-face dialogue) has many elements that are less explicit and highly context-dependent. Clark and Gerrig's (1990) description (and Peirce's symbols) are conventional and therefore relatively less context-dependent. Demonstrations, on the other hand, are idiosyncratic and may therefore be relatively more context-dependent, that is, more likely to require an ongoing dialogue. Although this explanation also dichotomizes monologue and dialogue, it is testable. For example, it would predict that overhearers should be better able to understand explanations made in a monologue condition than those made in a dialogue condition because the former would be more explicit.
3. On the other hand, it may be that the mere presence of another person—even without interaction—is socially arousing, and this arousal may lead to demonstrations, which tend to be more vivid and expressive (albeit not necessarily more accurate) than conventional description (in Clark & Gerrig's [1990] sense). If so, then even an imagined listener (the sixth context in Table 1) should elicit a higher rate of demonstrations than a pure monologue.
4. Another explanation would preclude an imagined listener. Clark and Gerrig (1990) proposed explicitly that demonstrations “work by enabling others to experience what it is like to perceive the things depicted” (p. 765). This implies that another person must be present (not just imagined) and capable of experiencing the demonstration. Again, however, no dialogic interaction would be required. For example, if this explanation holds, then when the speaker knows there is an audience who can see and hear him or her, but the speaker cannot see, hear, or interact with the audience, the rate of demonstrations should be similar to face-to-face dialogue.
5. Finally, because demonstrations are idiosyncratic, they may be more likely than description or indication (in Clark & Gerrig's [1990] sense) to require evidence of understanding in the moment-by-moment interaction with an addressee, that is, to require grounding. In a dialogue, the addressee can indicate when they do not understand the demonstration, and the speaker can clarify. This explanation requires a fully interacting addressee—without one, no such evidence would be available, so the rate of demonstrations should decrease.

We re-emphasize that these possible explanations are not mutually exclusive, that all have some initial plausibility, and that all would need more evidence. Our group's current emphasis is on the last possibility—that demonstrations depend on the potential for grounding that is uniquely available in dialogic

interaction. We are testing this possibility by comparing the mere presence of an interacting addressee with an addressee who is present but unable to interact with the speaker.

Importance of Dialogue as a Variable in Language Research

Finally, the current results draw attention to the broader methodological and theoretical importance of a dialogue–monologue continuum. The monologue condition in our experiments was similar to many experimental settings that focus on the language use of lone individuals, that is, on nonsocial speech. For example, studies of facial expressions, figurative language, and direct quotations often take place in contexts similar to our monologue condition. As Pickering and Garrod (2004) pointed out, the results from monologue experiments may be highly informative about individual processes, but they do not automatically generalize to language use in dialogue. For example, such results may greatly underestimate the rates at which speakers ordinarily use these demonstrations.

There is also increasing evidence that the results of studies with a confederate or the experimenter as addressee (i.e., the third context) do not necessarily generalize to free dialogues, as shown in the gesture visibility studies cited here. Lockridge and Brennan (2002), Kuhlen and Brennan (2013), and Bavelas and Healing (2013) raised questions about the use of confederates in several areas of language use. In any case, whenever the implicit purpose of a study is generalization to free dialogue, then including such a condition would enhance external validity. More broadly, using free dialogue versus monologue as an independent variable has revealed a novel and surprising effect on the rate of demonstrations. Introducing free dialogues into other research on language use may also lead to new insights.

ACKNOWLEDGMENTS

Members of our team who contributed in planning, conducting experiments, or analysis included Agustin Del Vento, Danielle Prevost, Chantelle Sutton, Christine Tomori, Patricia Wallis, and Jon Woods.

FUNDING

The Social Sciences and Humanities Research Council of Canada supported this program of research on face-to-face dialogue through grants to the first author.

REFERENCES

- Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*, 169–188.
- Bangerter, A., Mayor, E., & Pekarek Doehler, S. (2011). Reported speech in conversational storytelling during nursing shift handover meetings. *Discourse Processes*, *48*, 183–214.
- Bavelas, J. B., & Chovil, N. (1997). Faces in dialogue. In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression* (pp. 334–346). Cambridge, UK: Cambridge University Press.
- Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, *15*, 469–489.
- Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, *79*, 941–952.
- Bavelas, J. B., Gerwing, J., & Healing, S. (2014). Hand and facial gestures in conversational interaction. In T. Holtgraves (Ed.), *Handbook of language and social psychology* (pp. 111–130). Oxford, UK: Oxford University Press.
- Bavelas, J. B., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, *58*, 495–520.
- Bavelas, J. B., & Healing, S. (2013). Reconciling the effects of mutual visibility on gesturing: A review. *Gesture*, *13*, 63–92.
- Bavelas, J. B., Kenwood, C., Johnson, T., & Phillips, B. (2002). An experimental study of when and how speakers use gestures to communicate. *Gesture*, *2*, 1–17.
- Boerger, M. A. (2005). Variations in figurative language use as a function of mode of communication. *Journal of Psycholinguistic Research*, *34*, 31–49.
- Brugman, H., & Russel, A. (2004). Annotating multimedia/multimodal resources with ELAN. In M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa, & R. Silva (Eds.), *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004)* (pp. 2065–2068). Paris, France: European Language Resources Association.
- Brunner, L. J. (1979). Smiles can be back channels. *Journal of Personality and Social Psychology*, *37*, 728–734.
- Chovil, N. (1989). *Communicative functions of facial displays in conversation*. Unpublished Ph.D. dissertation, University of Victoria, Victoria, BC, Canada.
- Chovil, N. (1991). Social determinants of facial displays. *Journal of Nonverbal Behavior*, *15*, 141–153.
- Chovil, N. (1991/1992). Discourse-oriented facial displays in conversation. *Research on Language and Social Interaction*, *25*, 163–194.
- Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.
- Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, *66*, 764–805.
- Cohen, A. A. (1977). The communicative functions of hand illustrators. *Journal of Communication*, *27*, 54–63.
- Cohen, A. A., & Harrison, R. P. (1972). Intentionality in the use of hand illustrators in face-to-face communication situations. *Journal of Personality and Social Psychology*, *28*, 276–279.
- Colston, H. L., & Katz, A. N. (2005). *Figurative language comprehension: Social and cultural influences*. Mahwah, NJ: Erlbaum.
- Corts, D. P. (2006). Factors characterizing bursts of figurative language and gesture in college lectures. *Discourse Studies*, *8*, 211–233.
- Crystal, D. (2001). *A dictionary of language* (2nd ed.). Chicago, IL: University of Chicago Press.
- De Ruiter, J. P., Bangerter, A., & Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis. *Topics in Cognitive Science*, *4*, 232–248.

- Ekman, P. (1979). About brows: Emotional and conversational signals. In M. von Cranach, K. Foppa, W. Lepenius, & D. Ploog (Eds.), *Human ethology* (pp. 169–249). Cambridge, UK: Cambridge University Press.
- Ekman, P. (1997). Should we call it expression or communication? *European Journal of Social Sciences*, 10, 333–359.
- Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1, 49–98.
- Emmorey, K., & Casey, S. (2001). Gesture, thought and spatial language. *Gesture*, 1, 35–50.
- Fox, B. A., & Robles, J. (2010). It's like *mmm*: Enactments with *it's like*. *Discourse Studies*, 12, 715–738.
- Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiality by an implicit audience. *Journal of Personality and Social Psychology*, 60, 229–240.
- Fridlund, A. J., Sabini, J. P., Hedlund, L. E., Schaut, J. A., Shenker, J. I., & Knauer, M. J. (1990). Audience effects on solitary faces during imagery: Displaying to the people in your head. *Journal of Nonverbal Behavior*, 14, 113–137.
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27, 181–218.
- Garrod, S., & Doherty, G. (1994). Conversation, co-ordination, and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition*, 53, 181–215.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Science*, 8, 8–11.
- Gerwing, J., & Allison, M. (2011). The flexible semantic integration of gestures and words: Comparing face-to-face and telephone dialogues. *Gesture*, 11, 308–329.
- Glucksberg, S., & Krauss, R. (1967). What do people say after they have learned how to talk? Studies in the development of referential communication. *Merrill-Palmer Quarterly*, 13, 309–316.
- Holler, J., Tutton, M., & Wilkin, K. (2011). Co-speech gestures in the process of meaning coordination. *Proceedings of the 2nd GESPIN—Gesture in Speech and Interaction Conference*, Bielefeld, Germany. Retrieved from <http://gespin.amu.edu.pl/?q=node/66>
- Holler, J., & Wilkin, K. (2011). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*, 35, 133–153.
- Holsanova, J. (2006). Quotations as a vehicle for social positioning. In H. Hausendorf & A. Bora (Eds.), *Analysing citizenship talk: Discourse approaches to politics, society and culture* (pp. 251–275). Amsterdam, The Netherlands: Benjamins.
- Holsanova, J., Johansson, R., & Holmqvist, K. (2008). To tell and to show: The interplay of language and visualizations in communication. In P. Gardenfors & A. Wallin (Eds.), *A smorgasbord of cognitive science* (pp. 215–229). Nora, Sweden: Bokforlaget NyaDoxa.
- Koike, C. (2001). An analysis of shifts in participation roles in Japanese storytelling in terms of prosody, gaze, and body movements. *Berkeley Linguistics Society: Proceedings of the Annual Meeting*, 27, 381–392.
- Krauss, R. M., Dushay, R. A., Chen, Y., & Rauscher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology*, 31, 533–552.
- Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology*, 61, 743–745.
- Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review*, 20, 54–72.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Linell, P. (2005). *The written language bias in linguistics: Its nature, origins and transformations*. London, UK: Routledge.

- Linell, P. (2009). *Rethinking language, mind and world dialogically: Interactional and contextual theories of human sense-making*. Charlotte, NC: Information Age Pub.
- Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, 9, 550–557.
- Luckmann, T. (1990). Social communication, dialogue, and conversation. In I. Marková & K. Foppa (Eds.), *The dynamics of dialogue* (pp. 45–61). Hemel Hempstead, UK: Harvester Wheatsheaf.
- Marková, I. (1990). Introduction. In I. Marková & K. Foppa (Eds.), *The dynamics of dialogue* (pp. 45–61). Hemel Hempstead, UK: Harvester Wheatsheaf.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- McNeill, D. (2005). *Gesture and thought*. Chicago, IL: University of Chicago Press.
- Mol, L., Krahmer, E., Maes, A., & Swerts, M. (2009a). Communicative gestures and memory load. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1569–1574). Retrieved from <http://csjarchive.cogsci.rpi.edu/proceedings/2009/papers/339/index.html>
- Mol, L., Krahmer, E., Maes, A., & Swerts, M. (2009b). The communicative import of gestures. Evidence from a comparative analysis of human-human and human-machine interactions. *Gesture*, 9, 97–126.
- Peirce, C. S. (1940). *The philosophy of Peirce: Selected writings* (J. Buchler, ed.). London, UK: Routledge & Kegan Paul.
- Peirce, C. S. (1960). *Collected papers of Charles Sander Peirce, Volume IV: The simplest mathematics* (C. Hartshorne & P. Weiss, eds.). Cambridge, MA: The Belknap Press of Harvard University Press.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–226.
- Pine, K. J., Burney, D. J., & Fletcher, B. (2010). The semantic specificity hypothesis: When gestures do not depend upon the presence of a listener. *Journal of Nonverbal Behavior*, 34, 169–178.
- Rimé, B. (1982). The elimination of visible behavior from social interactions: Effects on verbal, nonverbal and interpersonal variables. *European Journal of Social Psychology*, 12, 113–129.
- Sams, J. (2007). Quoting the unspoken: An analysis of quotations in spoken discourse. *Colorado Research in Linguistics*, 20, 1–16.
- Sidnell, J. (2006). Coordinating gesture, talk and gaze in reenactments. *Research on Language and Social Interaction*, 39, 377–409.
- Wittenburg, P., Brugman, H., Russel, A., Klassman, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. *Proceedings of the Fifth International conference on language resources and evaluation*, Genoa, Italy. Retrieved from <http://www.lrec-conf.org/proceedings/lrec2006/>
- Yule, G. (1997). *Referential communication tasks*. New York, NY: Routledge, Taylor and Francis.

APPENDIX A

Sheet of Additional Instructions for Experiment 1

These instructions were the same for all conditions:

The following points are only intended to assist you in describing the excerpts of the movie that you will watch. Please address them in your general description of the excerpt, but include other parts of the movie as well. That is, do not restrict your description to these points. Read all of the points before watching the excerpts. Please watch the excerpts twice in order to remember as many details as possible.

(You do not have to rewind the tape; there are two copies of the clip, one after the other.) Please leave this card in this room when you are done. Thanks.

From the scenes in the first excerpt:

- Describe the scene when the cat encounters Shrek and Donkey.
- Describe Shrek’s reaction to the cat attacking him.
- Describe what Donkey does to help Shrek out and Shrek’s response to this.
- Describe what stops the cat from continuing to attack Shrek.

From the scenes in the second excerpt:

- Describe the cat’s strategy to persuade Shrek to take him along.

APPENDIX B

Events in *Shrek 2* Most Described by Speakers in All Conditions

<i>Event Number</i>	<i>Event Description</i>	<i>Analyzed for Facial Portrayals</i>	<i>Analyzed for Direct Quotations</i>
4	Shrek notices the cat (Puss in Boots) and says he is cute.		X
7	The cat scratches Shrek all over his body.		X ^a
8	Shrek screams for help and tries to get the cat off him.	X	X
11	Donkey aims for the cat and kicks Shrek instead.		X ^a
18	The cat falls to the ground, coughing up a hairball.		X ^a
25	The cat begs for mercy (with Spanish accent).		X
29	The cat stops Shrek from leaving and explains that he owes Shrek his life.		X
31	Donkey doesn’t want to take the cat along.	X	X
33	The cat begs silently, with big eyes.	X	X ^a
34	Shrek says that the cat is cute and wants to take him.	X	X

^a These events were excluded after analysis because virtually no quotations were located (see text).