

# You Said It Before and You'll Say It Again: Expectations of Consistency in Communication

Hadas Shintel and Boaz Keysar  
University of Chicago

Repeated reference creates strong expectations in addressees that a speaker will continue to use the same expression for the same object. The authors investigate the root reason for these expectations by comparing a cooperativeness-based account (Grice, 1975) with a simpler consistency-based account. In two eye-tracking experiments, the authors investigated the expectations underlying the effect of precedents on comprehension. The authors show that listeners expect speakers to be consistent in their use of expressions even when these expectations cannot be motivated by the assumption of cooperativeness. The authors conclude that though this phenomenon seems to be motivated by cooperativeness, listeners' expectation that speakers be consistent in their use of expressions is governed by a general expectation of consistency.

*Keywords:* referring expressions, language comprehension, lexical entrainment, mutual knowledge

Conversation is a ubiquitous human activity that seems trivially easy. However, studying the psychology of conversation has proven relatively difficult. A most influential proposal by Grice (1975) has reshaped current thinking about the logic behind a conversational exchange, explaining it as an inherently cooperative endeavor in which conversational partners work together to ensure mutual understanding. Since then, the assumption that people observe the cooperative principle, and that they assume that their interlocutors do the same, has made it easier to understand a variety of phenomena. Yet, recent research has begun to discover that some phenomena that seem to be motivated by an assumption of cooperative behavior are actually motivated by simpler processes (e.g., Bard et al., 2000; Barr, 2004a, 2004b; Brown & Dell, 1987; Keysar, 1997; Pickering & Garrod, 2004). We focus here on the following comprehension phenomenon that seems to be governed by the assumption of cooperativeness: When people converse, they establish a specific way of referring to things and continue to use the same expressions throughout the conversation. Most important, people expect their conversation partners to continue using the same expressions in the future.

In this article, we investigate whether this phenomenon is motivated by the assumption of cooperativeness or by a simpler expectation that is independent of Grice's principle of cooperative behavior.

## The Role of Precedents in Conversation

In everyday conversation, we continually refer to objects and easily understand such referring expressions. Communicating abstract ideas may sometimes pose difficulty, but referring to concrete objects seems relatively unproblematic. What could be more straightforward than understanding descriptions such as "the red handbag"? However, the apparent ease with which people communicate often masks the ambiguity inherent in referential communication. Even simple acts of reference involve a many-to-many mapping between referring expressions and their referents: A description such as "the red handbag" could pick out different handbags as referents, and the same handbag could be referred to by other descriptions such as "the leather purse." Successful communication requires the addressee to resolve the ambiguity and pick out the intended referent.

A central question in the field has been what strategies do people use to constrain the comprehension process and resolve referential ambiguity? One factor that past research has identified is lexical entrainment. With repeated reference, interlocutors entrain on specific referring expressions in the sense that they continue to use these same expressions in subsequent acts of reference (Brennan & Clark, 1996; Garrod & Anderson, 1987). Previous evidence has suggested that speakers continue using the same expressions when interacting with the same addressees, even in contexts in which these expressions are overinformative (Brennan & Clark, 1996), and that addressees indeed expect speakers to do so (Barr & Keysar 2002). By entraining on the same expressions during conversation, interlocutors can reduce referential ambiguity because they rely on an established mapping between an expression and a referent. Indeed, addressees are quicker to identify referents when they can rely on a precedent. Such an established word-referent mapping facilitates comprehension, especially when the referent lacks a conventional name and requires a somewhat unique description, which may be harder to interpret initially. For example when a speaker establishes the name *tent* for a folded, inverted V-shaped piece of paper, addressees' comprehension of

---

Hadas Shintel and Boaz Keysar, Department of Psychology, University of Chicago.

This work was supported by Public Health Service Grant R01 MH49685-06A1 to the University of Chicago. We would like to thank Kira Gole, Erica Kees, Brooke O'Neill, Abbe Rose, Anastasia Shakotko, Ashley Swanson, and Susan Swanson for their technical assistance with the experiments and Anne Henly for comments on earlier versions of this article. We also thank Simon Garrod for his criticism and suggestions.

Correspondence concerning this article should be addressed to Hadas Shintel, Department of Psychology, University of Chicago, 5848 South University Avenue, Chicago, IL 60637. E-mail: hadas@uchicago.edu

future reference to *tent* is facilitated (Barr & Keysar, 2002). Conversely, when speakers happen to violate their own precedents and use new referring expressions, addressees are delayed in identifying the intended referents and are more likely to initially interpret the new expressions as referring to a different object (Keysar, Barr, & Lim, 2001; Metzling & Brennan, 2003).

The cost associated with the violation of conversational precedents may reflect listeners' expectation that speakers adhere to their linguistic precedents. As indicated by Metzling and Brennan (2003), this pragmatic expectation is predicted by E. V. Clark's (1988, 1990) principle of contrast, which states that every two forms must contrast in meaning. Underlying the principle of contrast is the intuition that by choosing a particular referring expression, speakers adopt a particular perspective or a way of conceptualizing the object. When a speaker uses a different linguistic form, it can be taken to indicate a difference in underlying intention. From this point of view, a change in form can signal a change in intended reference by way of a pragmatic inference: Had the speaker wanted to refer to the same object, the speaker would have used the same referring expression. Of course, a change in form may indicate a change in sense with no change in reference. For example you may say "that old thing" to refer to the same object you previously called "the silk shirt," thereby indicating a change in the way you now view it. However, in the absence of a reason motivating a change, speakers are expected to continue using the same terms. This kind of pragmatic account was also proposed for children's avoidance of lexical overlap (Diesendruck & Markson, 2001; see also Bloom, 2000). When presented with a novel object and a familiar object, 3-year-olds tend to choose the novel object as the referent of a novel name. This phenomenon may reflect the same pragmatic constraints underlying adults' reaction to a violation of a precedent, rather than a purely lexical constraint (cf. E. V. Clark, 1990; Gathercole, 1989; for a contrasting view, see Markman, 1990).

One important prediction follows from this account: Listeners' expectations should be directed at the specific speaker who established the precedents and should not generalize to other conversational partners. In contrast, if a conversational precedent affects comprehension only because of the availability of an established word-referent mapping, its effect should be independent of the identity of the speaker (Barr & Keysar, 2002). Indeed, the cost associated with the violation of a precedent appears to be speaker specific (Metzling & Brennan, 2003). In Metzling and Brennan's study, comprehension was delayed when speakers used new expressions that violated their own precedents but not when new speakers used the same new expressions. Furthermore, comprehension was delayed when speakers violated their own precedent, even when the new expression had been previously used by a different speaker to refer to the same object, thereby establishing a mapping of that expression to the referent (Keysar et al., 2001). This pattern of results suggests a speaker-specific expectation. Listeners may interpret a change in form as implicating a change in referent but only when this change occurs within a speaker and not between speakers.

Although the comprehension benefit associated with a precedent appears to be independent of whether the original speaker or a new speaker uses the expression (Barr & Keysar, 2002), such a speaker-independent benefit may be consistent with speaker-specific expectations: The repetition of an expression may facili-

tate comprehension so powerfully that it may swamp the potential effect of a partner-specific cue. Therefore, it may be more productive to look for speaker specificity where it is more likely to be found, when speakers violate their own precedents (Metzling & Brennan, 2003).

### Partner Specificity Versus Speaker Specificity

Here we are interested in further investigating this speaker-specific interference effect and the situations in which it comes into play. We consider two accounts, a cooperativeness-based account and a general consistency-based account. According to a cooperativeness account, the speaker-specific effect may reflect listeners' expectations about the nature of communication. If speakers are cooperative (Grice, 1975), they should make their utterances as unambiguous as possible and, hence, should rely on existing precedents that they had previously established with their interlocutors. This kind of explanation underlies the suggestion by Brennan and Clark (1996; see also Metzling & Brennan, 2003) that both lexical entrainment and partner-specific effects in comprehension result from creating, or breaking, a partner-specific conceptual pact: By calling a shoe a loafer, the partners create a pact to conceptualize the specific shoe in a particular way. Such a pact entails an expectation that the conversational partner will adhere to the precedent in future interaction with the same partner because they, and not others, made this pact. Critically, the cooperativeness account suggests that interlocutors' beliefs about the mutuality of a precedent underlie the speaker-specific effects of the violation of linguistic precedents. A change in expression implicates a change in referent only if the addressee believes that in choosing a different expression the speaker has broken an established mutually known conceptual pact and indeed intends to signal such a change (see Brennan & Metzling, 2004; Grice, 1989). The addressee therefore needs to consider the common ground, specifically the conceptual pacts previously established with the specific speaker.

According to a consistency account, speaker specificity may arise from a more general assumption that speakers will be consistent in the absence of a discernible reason for inconsistency. Consistency, on this account, is not addressee directed and does not entail cooperativeness. Listeners simply use their knowledge about speakers' previous behavior in interpreting their present behavior or predicting their future behavior. Speaker specificity does not require the listener to model the speaker's beliefs and intentions. On this account, the cost associated with the violation of precedents is indeed speaker specific, but it is not partner specific: Listeners may expect speakers to adhere to their own precedents independently of the conversational partner.

If listeners expect speakers to be consistent because they expect them to be cooperative, this expectation should hold only in situations in which the precedent can be assumed to be mutually known. On the other hand, if listeners' expectation of consistency arises simply from inferences based on speakers' past behavior, it should hold independently of whether the precedent is mutually known. On this account, listeners' expectation will hold even if they believe that speakers do not know they share this knowledge, and hence, this knowledge cannot be assumed to be mutual. Accordingly, this expectation is not dependent on the listener being part of a conceptual pact regarding the use of the expression.

The idea that repeated use of a precedent will result in a speaker-independent benefit is consistent with both of these accounts. Neither an expectation of cooperativeness nor a general expectation of consistency implies that listeners should be surprised if different speakers use the same expression or if a speaker uses the same expression when interacting with different addressees. However, these accounts make different predictions regarding the interference due to a violation of precedents. Although both are consistent with a speaker-specific violation effect, only the cooperativeness account further predicts that such an effect is partner specific (that is, addressee specific in addition to speaker specific). Because repeated use reflects accommodation to the addressee's needs, a violation should affect comprehension only when the precedent is part of the mutual knowledge the addressee has with the speaker. Thus the cooperativeness account predicts interference only when a speaker violates a precedent that was previously established with that specific addressee. On the other hand, a consistency account predicts a violation effect even if the speaker violates a precedent that was never established with that addressee but was established with a different addressee.

The evidence obtained so far in support of speaker-specific effects (Keysar et al., 2001; Metzger & Brennan, 2003) focused on addressees' expectations when interacting with different speakers. However, these findings did not examine addressees' expectations when the same speaker interacts with different addressees. Consequently, these findings cannot discriminate between speaker specificity and partner specificity, and thus they cannot discriminate between the two explanations. To evaluate these possibilities, we conducted two experiments in which we manipulated addressees' beliefs about the mutuality of the precedents. If speaker-specific effects are due to an expectation of cooperativeness, they should be evident only when the precedent is presumed to be mutually known by speaker and addressee. On the other hand, if speaker-specific effects are governed by a general expectation of consistency, independently of an assumption of cooperativeness, they should be evident even when the precedent is not mutual.

### Experiment 1

Consider the following situation. You are helping your friend John and his roommate Mark to pack before moving. At a certain point John asks Mark to hand him "the blue picture," and Mark gives him a painting of a city skyline that has a lot of blue sky in it. The next day, only John and you are in the room and John asks you for "the city picture." Consistent with published research, because of the change in form you would probably hesitate temporarily before you would determine that the skyline picture is indeed the intended referent. If you hesitate because you expect John to be cooperative, this hesitation depends crucially on your belief that John knows that you are familiar with the precedent, because you were in the room when he referred to the picture as "the blue picture." Now consider a slightly different scenario. You observed John asking Mark for "the blue picture" but you know that John thought you were away at the time. In this situation, John does not know that you are familiar with the referring precedent. Therefore, if on the next day John asks you for "the city picture," you should not infer that he is intending to refer to a different picture: As far as John knows, you are not familiar with the precedent (i.e., "the blue picture") and therefore would not per-

ceive the new expression as a change in form. So, according to the cooperativeness account, perceiving the change should lead to hesitation only when you believe that John knows you are familiar with the precedent, that is only when the belief that you are familiar with the precedent is part of your common ground. In contrast, the general consistency account predicts that your expectation that John would continue to use the same term would be independent of such mutuality. You will hesitate regardless of whether John thinks you are familiar with the precedent. The logic of Experiment 1 follows the logic of this scenario.

In this experiment, speaker and addressee pairs participated in a referential communication task in which a female confederate participant (the director) instructed the participant (the matcher) to rearrange objects in a grid to match a picture. Before the test trials began, matchers watched videos of the director performing the same task with another matcher. We used the video as the means of informing the participant matcher about the precedents that the director established in the past. For example, in the video the director referred to an elephant-shaped baby rattle as "the elephant rattle." After the matcher watched the video, the director instructed him or her and referred to objects either with the expression from the video (e.g., "Move the elephant rattle") or with a new expression, such as "Move the baby rattle." See Figure 1.

The critical manipulation involved matchers' beliefs about the director's knowledge. In one condition (the knowledge condition), matchers watched the videos along with the director. Therefore, the referring expressions the director used in the video can be viewed as part of the common ground of the director and the matcher. In the second condition (the no-knowledge condition), the director was not present when the matchers watched the video. In addition, the matchers were led to believe that the director was not aware that they had watched the video. Therefore, as far as the matchers knew, the director had no reason to believe that they knew what terms the director used in the video.

We tracked matchers' eye gaze as they heard the instructions, as an index of the degree to which they considered an object as a potential referent at different stages of the comprehension process (Cooper, 1974; Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995). Specifically, matchers' eye gaze served as an index of the degree to which matchers considered an object as a potential referent even prior to making their final decision. If addressees expect speakers to use the same terms because of cooperativeness, they should be delayed in identifying intended referents when the director violates her precedent only where precedents are mutually known—that is, in the knowledge condition but not in the no-knowledge condition. The cooperativeness account would therefore predict an interaction between the violation effect and the knowledge condition. On the other hand, if addressees expect speakers to be consistent independently of cooperativeness but as part of a general expectation of consistency, they should be similarly delayed whenever the speaker is inconsistent in her use of terms, both in the knowledge and in the no-knowledge conditions.

### Method

*Participants.* Thirty-six University of Chicago college students participated in the study for payment. Data from 3 additional participants were not usable because of calibration problems or experimenter's errors.

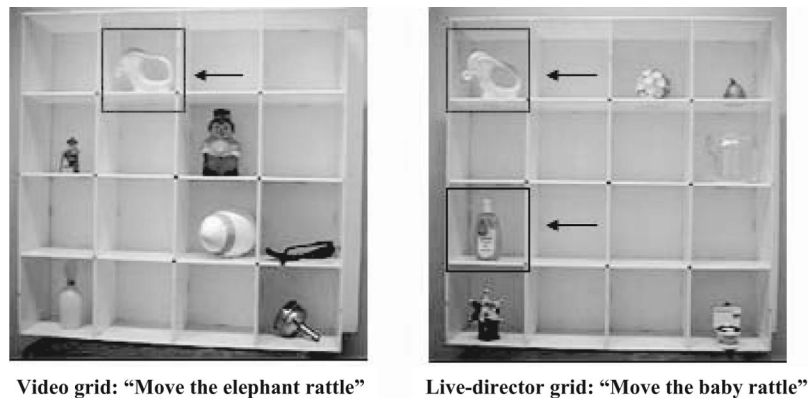


Figure 1. Picture of the video grid and the live director grid for one experimental item in Experiment 1.

**Apparatus.** The grid was composed of 16 boxes arranged in a  $4 \times 4$  pattern. Each square in the grid measured  $12.5 \times 12.5 \times 12.5$  cm. Objects were visible from both sides of the grid. We monitored participants' eye movements with an iView X head-mounted eye-tracking system (SensoMotoric Instruments, Teltow, Germany). Head-mounted cameras captured the matcher's eye movement and the scene from the matcher's point of view. A gaze cursor indicating the computed gaze position was overlaid on the scene image. Overlays were recorded as MPEG videos at a temporal resolution of 30 Hz (approximately one frame every 33 ms). A PC running iView X software (SensoMotoric Instruments [SMI], Teltow, Germany) digitally stored the real value coordinates of the matcher's gaze at a rate of 60 Hz. A microphone, placed on the director's side of the desk, recorded the speakers' instructions onto the MPEG video.

**Materials and design.** Critical objects were placed in the grid along with filler objects. There were five sets of objects for each stage of the experiment: five video grids and five live-director test grids. The director gave eight to nine instructions for each grid. For the video, we used a digital video camcorder to film the director instructing a confederate matcher. Videos were filmed from the matcher's point of view; both the director and the grid were visible in the videos. Each critical object was mentioned three or four times, either as the target (e.g. "Take the elephant rattle and put it above the clown") or as a landmark (e.g. "Put the flower above the elephant rattle"). Videos were presented on a computer screen in a fixed order.

In each live-director test grid, the director gave two critical instructions and six to seven filler instructions. The critical instructions mentioned an object previously referred to in the videos by using either the same expression as in the video or a new expression. See the Appendix for a list of the original and new expressions. The critical instruction was the first time the target object was referred to in the live-director test grids. In addition to the target objects, each grid contained an object that provided an opportunity for temporary ambiguity of the new description. For example if the new description was "the baby rattle," one of the objects was a small bottle of baby soap. By the second word, the description uniquely identified the target object (see Figure 1). There was no other object in the display that could have been described with the same noun as the potential competitor. So the

situation could not have invoked a referential contrast that would have led matchers to interpret the adjective as referring to the potential competitor (Sedivy, 2003). For example, the display contained no other soap that could lead matchers to interpret "baby" as indicating a contrast between the baby soap and the other kind of soap. Furthermore, no competitor object could have been referred to with either the adjective or the noun of the old expression; the different context of the test grid did not render the old expressions underinformative so as not to introduce a motivation for changing the expression on the part of the speaker. In fact, whereas new expressions were temporarily ambiguous, old expressions were not. Live-director grids were presented in a random order. The competitor was never referred to by the director. Items were counterbalanced across participants, creating two versions of the instructions in the actual task phase. The configuration of each grid was the same for all participants.

The design was a  $2$  (director knowledge: knowledge vs. no-knowledge)  $\times 2$  (expression: old vs. new) randomized block design with director knowledge as a between-subjects factor and expression as a within-subjects factor. Participants were randomly assigned to conditions with the purpose of having an equal number of participants in each cell. A comparison between the old and the new expression conditions involves a comparison across different expressions and therefore should be carefully interpreted. Therefore, the critical prediction concerns the Knowledge  $\times$  Expression interaction; such an interaction is predicted by the cooperativeness account, whereas the consistency account predicts a similar effect of expression across the two knowledge conditions.

**Procedure.** To ensure that participants heard the same critical instructions, we trained a female confederate as the director in the experiment to give instructions in a seemingly spontaneous way. After delivering the scripted instruction, the director spontaneously responded to the matcher's questions. Participants were led to believe that they were interacting with a naive participant. In both conditions (and in the videos), the director wore dark sunglasses to prevent matchers from following her eye gaze while she gave the instructions.

In the knowledge condition, participants arrived at the same time at the lab. The naive participant was told that the director had completed the same task the day before with another matcher and that for practice they are going to watch the videos from that part



of the study. Participants watched one filler video together. For the five critical video grids, the matcher sat in front of the grid and was asked to actually carry out the director's instructions. Throughout the whole time, the director watched the videos in the same room. The sound was played through speakers so both the director and the matcher could hear the instructions. After completing the five grids, matchers were fitted with the eye tracker and were calibrated, and then they completed five grids with the live director.

In the no-knowledge condition, matchers were told they were about to perform a referential communication task with another participant who had completed the task with another matcher the day before. The experimenter informed them that the director was scheduled to arrive later and that, for practice, they were going to watch the videos of the task the director had performed the day before. As in the knowledge condition, the matchers carried out the instructions from the video. After finishing the video grids, they were told that only half of the matchers watch the videos and as all other aspects of the experiment should remain constant, directors should not know which matchers watched the videos. Matchers were asked not to mention watching the videos and to pretend that they had just arrived. When the director arrived, matchers went through the calibration procedure and proceeded to complete five grids with the live director.

After completing the task, all participants received a postexperimental questionnaire designed to check whether they guessed that the director was a confederate and whether they suspected she knew that they had watched the videos. The postexperimental questionnaire indicated that in general the cover story was effective and that participants believed that the director in the second game was a naive participant. Out of 18 participants in the no-knowledge condition, 2 thought the director was trained beforehand and considered this possibility during the experiment. Only 1 of them believed the director actually knew or suspected he had watched the videos. This participant was excluded from the analysis. Therefore, participants' performance in the no-knowledge condition could not have resulted from a belief that the director was aware that they had knowledge about the expressions used in the videos. Four out of 18 participants in the knowledge condition reported guessing the director was trained beforehand and indicated that they considered this possibility during the experiment. None of these participants guessed the purpose of the experiment or reported noticing the director used different expressions to refer to the same object. After completing the questionnaire, participants were debriefed about the purpose of the study and the reason for the deception.

*Coding and analysis.* A college student, unfamiliar with the hypothesis of the study and blind to the different conditions, coded the digital video data files. For each critical utterance, the coder located the *onset* of the target word, defined as the onset of the initial syllable of the adjective at the beginning of the description pronounced by the director, and the *end point*, defined as the point of touching the object. These two points determined the window of observation for each trial. SMI iView X analysis software was used to extract the data for the coordinates for the location of the eye at each sample from the digital data. These data were used to compute the fixations on the critical objects. The criterion for a fixation was that the eye gaze remained on the critical square in the grid for at least 100 ms, about three consecutive frames.

## Results and Discussion

To prevent inflation because of clarification or confirmation requests, response times were truncated separately for each condition to a value of 2.5 standard deviations from the mean of the distribution for the trials in which there were no clarification or confirmation requests. The truncation procedure affected 5% of the data for the latency to the first fixation and the touch time, distributed evenly between the knowledge and no-knowledge condition. The pattern of results was the same before and after the truncation procedure. Complete means and standard errors for all measures appear in Table 1.

The first fixation on the target object is an index of the initial noticing of the target. There was a main effect of expression,  $F_1(1, 33) = 35.18$ ,  $MSE = 6,579,178$ ,  $p < .0001$ ;  $F_2(1, 9) = 23.40$ ,  $MSE = 1,998,655$ ,  $p < .001$ . In general, matchers took longer to fixate on the target object when the director used a new expression compared with an old one ( $M_s = 1,497$  ms and 884 ms, respectively; see Figure 2). Given that the two expression conditions used different utterances, these latencies cannot be directly compared. However, it is important to note that reaction times for each expression condition were similar across the two knowledge conditions, as shown by the lack of a Knowledge  $\times$  Expression interaction ( $F_s < 1$ ). So in contrast to the speaker-specific hypothesis, there was no evidence that the delay was larger in the knowledge condition.<sup>1</sup> Results suggest that the delay for new expressions was independent of director knowledge, supporting the general consistency hypothesis.

The time to touch the target is an index of the final decision on the referent. These results mirror the results for the first fixation: There was no evidence of interaction between director knowledge and expression ( $F_s < 1$ ). The main effect of expression was significant,  $F_1(1, 33) = 28.24$ ,  $MSE = 5,443,614$ ,  $p < .0001$ ;  $F_2(1, 9) = 7.60$ ,  $MSE = 3,251,355$ ,  $p < .03$ . Director knowledge was not significant in the participants analysis or in the item analysis,  $F_1 < 1$ ;  $F_2(1, 9) = 3.93$ ,  $MSE = 128,843$ ,  $p = .10$ .

We also analyzed both the number of looks at the potential competitor and their duration, as an indicator that participants are considering alternative referents. Given that the competitor did not represent a good match for the old expressions, we expected matchers to look at it more in the new expression condition. Indeed, matchers looked at the competitor at least once when the director used a new expression, but not when she used the old expression, and fixations on the competitor were longer with new expressions. The main effect of expression was significant: number of fixations,  $F_1(1, 33) = 43.65$ ,  $MSE = 6.40$ ,  $p < .0001$ ;  $F_2(1, 9) = 6.83$ ,  $MSE = 3.67$ ,  $p < .03$ ; duration of fixations,  $F_1(1, 33) = 58.66$ ,  $MSE = 46,004,410$ ,  $p < .0001$ ;  $F_2(1, 9) = 17.30$ ,  $MSE = 4,504,173$ ,  $p < .003$ . If the violation-of-precedent interference is due to expectations of cooperativeness, then matchers should look more at the competitor in the knowledge than the no-knowledge condition, because the precedent is mutual only in the knowledge condition. However, there was no hint of a Director Knowledge  $\times$  Expression interaction ( $F_s < 1$ ). Matchers were just as likely to look at the competitor object when the director used a new ex-

<sup>1</sup> In fact, nominally the delay was bigger in the no-knowledge condition, demonstrating that the failure to find a greater delay in the knowledge condition is not due merely to lack of statistical power.

Table 1  
Means and Standard Errors by Knowledge and Expression in Experiment 1

Dependent variable	Knowledge				No knowledge			
	New expression		Old expression		New expression		Old expression	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
First fixation latency (ms)	1,494	112	959	68	1,500	176	808	58
Touch time (ms)	3,637	267	3,073	237	3,512	120	2,960	201
No. of fixations on competitor	1.13	0.12	0.49	0.08	1.07	0.10	0.50	0.06
Duration of fixations on competitor (ms)	948	176	263	103	732	110	220	32

pression in the no-knowledge condition as they were in the knowledge condition. The main effect of director knowledge was not significant ( $F_s < 1$ ).<sup>2</sup>

One possible concern is that because the confederate director was aware of the two different knowledge conditions, she might have said the critical expressions differently in the different conditions. Furthermore, the director heard the expressions used in the video in the knowledge condition but not in the no-knowledge condition; hearing the words before saying them may have caused her to reduce the duration and intelligibility of the referring expressions (Bard et al., 2000). To examine these possibilities, we analyzed the director's utterances in the two director knowledge conditions. A coder blind to the hypotheses and to the conditions measured the duration of each token of referring expression, the total duration of the critical instruction, and the number of hesitations during each critical instruction. Complete means are presented in Table 2. There was no significant difference between the two director knowledge conditions for any of the measures, nor was there a significant Director Knowledge  $\times$  Expression interaction ( $F_s < 1$ ). We did not compare the two expression conditions (old vs. new) that involved different referring expressions. Although we cannot conclusively determine that there was no relevant difference in the director's speech between the two conditions, these results strongly suggest that the pattern of results found for matchers' behavior does not result from differences in production of the critical referring expressions.

Our results suggest that when the speaker violated her own precedents, listeners were delayed in identifying the intended referent and initially interpreted the change in expression to mean a change in intended referent. These results replicate previous

findings that when speakers violate their own precedents, it results in a cost to comprehension (Keysar et al., 2001; Metzger & Brennan, 2003). More important, however, our results show that listeners were just as delayed in identifying the intended referents in the no-knowledge condition as they were in the knowledge condition. Listeners expected the director to consistently use referring expressions even when they believed she was unaware of their knowledge of past expressions. If listeners expect the speaker to continue using the same referring expression because of cooperativeness, they should have this expectation only when the speaker believes that they have knowledge of previous usage patterns, namely, only in the knowledge condition. In the no-knowledge condition, the assumption of cooperativeness does not motivate an expectation of consistency. These results show not only that matchers were clearly delayed in the no-knowledge condition but also that this delay was not smaller than the delay in the knowledge condition. Listeners, then, expect speakers to use the same expressions in conversation, independently of cooperativeness.

One potential problem with Experiment 1 concerns the nature of the knowledge manipulation and whether it was indeed sufficient for establishing mutuality in the knowledge condition. H. H. Clark and Marshall (1981) proposed that co-presence provides a basis for inferring mutuality only if it meets the standard of triple co-presence, according to which speaker, addressee, and referent are all "openly present together" (p. 32). Triple co-presence entails both the assumption of simultaneity (speaker and addressee are both looking at the referent and at each other) and the assumption of attention (speaker and addressee are both attending to the referent and to each other). In the knowledge condition, both the director and the matcher were present during the video presentation of the director's "first game." However, it is possible that the matcher was not completely aware of the fact that they were jointly attending to the referring expressions in the video. Because the

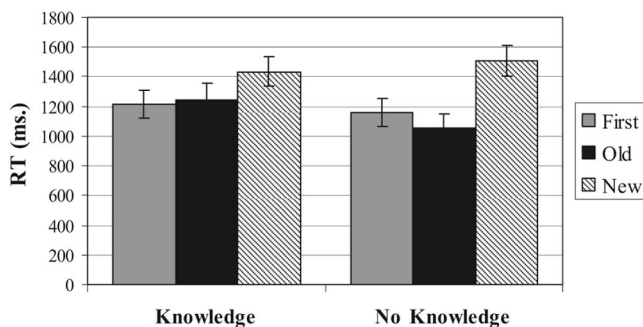


Figure 2. Latencies of first fixation on target for the expression-duration matched subsample of the data in Experiment 2. RT = response time.

<sup>2</sup> We also analyzed the results excluding the additional 5 participants (1 in the no-knowledge condition and 4 in the knowledge condition) who suspected the director was trained beforehand to ensure that this did not affect the results. The analysis showed the same pattern. There was a significant effect of expression on the first fixation to the target,  $F(1, 28) = 29.87$ ,  $MSE = 6,194,430$ ,  $p < .0001$ ; touch time,  $F(1, 28) = 23.52$ ,  $MSE = 4,181,913$ ,  $p < .0001$ ; and on looks to the competitor object, number of fixations,  $F(1, 28) = 42.52$ ,  $MSE = 5.80$ ,  $p < .0001$ ; duration of fixations,  $F(1, 28) = 52.25$ ,  $MSE = 4,956,242$ ,  $p < .0001$ . All other effects were not significant ( $p > .2$ ).

Table 2  
*Director's Utterances in Experiment 1*

Dependent variable	Knowledge				No knowledge			
	New expression		Old expression		New expression		Old expression	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Duration of expression (ms)	861	33	777	19	891	37	783	26
Total duration of utterance (ms)	5,361	321	4,169	91	5,279	262	4,318	117
No. of hesitations	1.93	0.18	1.38	0.11	1.95	0.19	1.37	0.11

matchers' task was to actually follow the director's instructions in the video, we believe matchers were attending to the referring expressions. Such attention is further implied by the switch cost during the test trials. However, it is possible that matchers were not able to establish that the director was indeed attending to her own referring expressions or to the fact that they were correctly understood by matchers. In this case, the matcher and the director may not be jointly attending to the director's referring expressions. Consequently, matchers would not necessarily expect cooperativeness to entail the director adhering to her own precedents. A related concern is that the practice phase in the experiment did not involve actual interaction between director and matcher, and thus the matcher was not acting as an addressee but essentially as an overhearer. As pointed out by Schober and Clark (1989), speakers are not responsible for making themselves understood to overhearers, and overhearers do not acknowledge the establishment of a mutual belief about the referent. Thus, the procedure may not have been powerful enough for establishing the referring expressions as mutually known linguistic precedents.

Although the triple co-presence constraint appears too strong as a general constraint, as there are many cases of establishment of common ground under weaker conditions, we conducted a second eye-tracking experiment to see whether our conclusions would hold even under such strict criteria. The critical manipulation in Experiment 2 was similar to Experiment 1; however, instead of having matchers watch a video in the director's presence, the knowledge condition involved an actual interaction between the director and the matcher, in which the matcher acted as an addressee and was able to ask for more information and to verify that the mapping between the expression and the referent had been grounded—that is, made part of the common ground between director and matcher. The no-knowledge condition remained essentially as in Experiment 1. If the pattern of results observed in Experiment 1 was due to the specific knowledge manipulation introduced in that experiment, we would expect a Knowledge  $\times$  Expression interaction to emerge in Experiment 2. If, however, the results still hold under the stricter conditions of Experiment 2, that would add support to the general consistency hypothesis.

Experiment 2 was also designed to address a potential difficulty in interpreting the findings of Experiment 1. The results showed that the violation of a precedent results in comparable interference in the two knowledge conditions, in contrast to the predictions of the cooperativeness account. We showed this by comparing latencies in the old and the new expression conditions. It is possible, though, that the latency for a new expression was longer not because it violated a precedent but because it is not as good a

match for the object compared with the old expression. To show that the expectation of consistency is independent of the mutuality of precedents, we have to establish the existence of such an expectation in our experimental setting by showing a clear cost for the violation of a precedent. To control for that, in Experiment 2 we added a baseline condition in which the new referring expressions were used without a precedent. If longer reaction times to new expressions indeed reflect interference due to the violation of a precedent, then latency with new expressions should be longer when a precedent exists.

## Experiment 2

As in Experiment 1, participant matchers participated in a referential communication task in which a confederate director instructed them to rearrange objects in a grid in order to match a picture. The experimental session consisted of two phases, a practice phase and the test phase. In the knowledge condition, matchers performed the same referential communication task with the director both in the practice phase and in the test phase. The no-knowledge condition essentially replicated the no-knowledge condition in Experiment 1: Matchers watched videos of the director instructing another matcher. The director was not present during the practice phase, and matchers were led to believe that she was not aware that they watched the videos. Thus, the expressions used in the video were known to both director and matcher, because the director used them and the matcher saw her use them in the video, but the matchers had no reason to believe these expressions were mutually known. The test phase of the experiment was identical across the two knowledge conditions. In this stage, matchers performed the task with the same director. If the comprehension cost associated with violation of precedents is dependent on the assumption of cooperativeness, such a cost should be bigger in the knowledge condition compared to the no-knowledge condition. Alternatively, if the violation cost depends on a more general expectation of speaker consistency, we would expect the same cost in both knowledge conditions.

## Method

**Participants.** Thirty-nine University of Chicago college students participated in the study for payment. Two additional participants did not provide usable data because of calibration problems or experimenter's errors.

**Apparatus.** The apparatus (a 4  $\times$  4 grid and an SMI iView X head-mounted eye-tracking system) was identical to the apparatus used in Experiment 1.

*Materials and design.* Critical objects were placed in the grid along with different filler objects. There were four sets of objects for each stage of the experiment: four practice (either live or video) grids and four test grids. A female confederate director gave eight to nine instructions for each grid. Video grids were filmed from the matcher's point of view with both the director and the grid visible.

Each practice grid contained three critical objects that were used as the target objects in the test phase of the experiment. One of these critical objects appeared in the practice grids but was not referred to by the director. Each of the other two critical objects was mentioned three or four times, either as the target (e.g., "Take the elephant rattle and put it above the clown") or as a landmark (e.g., "Put the flower above the elephant rattle"). Practice grids were presented in a random order.

In each test grid, the director gave three critical instructions and five to six filler instructions. There were three kinds of critical instructions (one of each kind in each of the four test grids), corresponding to three expression conditions. One critical instruction mentioned an object previously referred to in the practice phase by using the same expression (the old condition). A second critical instruction mentioned an object previously referred to in the practice phase by using a new expression (the new condition). Finally, a third critical instruction mentioned for the first time an object that was seen in the practice phase but was not referred to (the first condition). The expressions used in the first condition were the same as the expressions used in the new condition, except that no precedent was established in the practice phase. Each critical object appeared in all three expression conditions, counterbalanced across participants.

As in Experiment 1, in addition to the target objects, each test grid contained an object that provided an opportunity for temporary ambiguity of the new expression. Because the same expressions were used for the first time both in the first condition and in the new condition, the first condition can serve as a control condition for evaluating whether any potential delay in comprehension is due to the violation of a precedent rather than to the difficulty or the temporary ambiguity of the expression.

Test grids were presented in a random order. The design was a 2 (director knowledge: knowledge vs. no knowledge)  $\times$  3 (expression: old vs. new vs. first) randomized block design with director knowledge as a between-subjects factor and expression as a within-subjects factor. Items were counterbalanced across participants, creating three versions of the instructions in the actual task phase. Participants were randomly assigned to conditions with the purpose of having an equal number of participants in each cell.

*Procedure.* In the knowledge condition, both participants arrived at the same time. To make the two conditions as comparable as possible, the matcher was told that the director completed the same task the day before with another matcher. Participants watched one filler video together as a demonstration of the task. Participants then continued to complete the task for four practice grids. After completing the four grids, matchers were fitted with the eye-tracker and then completed the four test grids with the same director. In all other respects, the procedure was the same as in Experiment 1. The no-knowledge condition was identical to the no-knowledge condition in Experiment 1. The postexperimental questionnaire indicated that in general the cover story was effective and that participants believed that the director in the second game was a naive participant. Out of 20 participants in the no-

knowledge condition, four participants thought the director was trained beforehand and reported considering this possibility during the experiment. Only one of them believed the director actually knew or suspected she had watched the videos; this participant was excluded from the analysis. Four out of 19 participants in the knowledge condition thought the director was trained and reported considering this possibility during the experiment. None of these participants guessed the purpose of the experiment. Only one of them reported noticing the director referred to the same object using two expressions. That participant was excluded from the analysis. After completing the questionnaire, participants were debriefed about the purpose of the study and the reason for the deception.

*Coding and analysis.* Two college students, unfamiliar with the hypothesis of the study and blind to the different conditions, coded the digital video data files. The coding procedure was the same as in Experiment 1.

## Results and Discussion

Seven trials out of a total of 468 trials ( $< 2\%$ , each from a different participant) were not included in the analysis because of experimenter's errors or confederate director's errors on those trials. Response times were truncated to prevent inflation due to clarification questions or confirmation requests. Response times were truncated to a value of three standard deviations from the mean of the distribution for that condition. Response times greater than 10 s were not included in the computation of the mean and the standard deviation used for fixing the truncation values. The truncation procedure affected 4.3% of the data for the first fixation, distributed evenly between the knowledge and no-knowledge conditions, 3% of the data for the touch point (53% from the knowledge condition and 47% from the no-knowledge condition), and 3% of the data for the duration of fixations on the competitor object (36% in the knowledge condition and 64% in the no-knowledge condition). Complete means and standard errors for all measures appear in Table 3.

Data were subjected to a 2 (knowledge: knowledge vs. no-knowledge)  $\times$  3 (expression: first vs. old vs. new) analysis of variance with knowledge as a between-subjects factor and expression as a within-subjects factor. As expected, matchers were slower to touch the target object in the new condition, suggesting they took longer to decide on the target object as the intended referent. The effect of expression was significant both by participants and by items,  $F_1(2, 70) = 24.54$ ,  $MSE = 4,134,772$ ,  $p < .0001$ ;  $F_2(2, 22) = 8.40$ ,  $MSE = 2,473,106$ ,  $p < .005$ . Pairwise comparisons between the different expression conditions using the Bonferroni-Dunn correction (the significance level required for significance was .0167) revealed that the new expression condition differed significantly both from the old expression condition ( $ps < .002$ ) and from the first expression control condition ( $ps < .003$ ), suggesting that the delay in the new condition was not due merely to the temporal ambiguity of the referring expression but resulted from the violation of a precedent. The difference between the old and the first expression conditions was not significant ( $ps > .6$ ). The main effect of director knowledge was significant only in the items analysis,  $F_1(1, 35) = 2.47$ ,  $MSE = 6,641,814$ ,  $p > .1$ ;  $F_2(1, 11) = 33.65$ ,  $MSE = 4,923,496$ ,  $p = .0001$ . There was no significant Director Knowledge  $\times$  Expression interaction ( $F_s < 1$ ).



Table 3  
Means and Standard Errors by Knowledge and Expression in Experiment 2

Dependent variable	Knowledge						No knowledge					
	First		New		Old		First		New		Old	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
First fixation (ms)	1,228	83	1,466	92	1,227	100	1,114	69	1,504	88	1,001	90
Touch time (ms)	2,717	128	3,308	215	2,738	171	3,276	276	3,793	264	3,161	313
No. of fixations on competitor	0.61	0.09	0.74	0.12	0.32	0.05	0.49	0.09	0.78	0.08	0.38	0.07
Duration of fixations on competitor (ms)	208	35	311	51	113	23	170	35	410	65	112	26

A similar effect of expression was evident in the latencies of first fixation on the target object—that is, before matchers made their final decision regarding the intended referent. There was a significant effect of expression,  $F_1(2, 70) = 21.18$ ,  $MSE = 1,476,934$ ,  $p < .0001$ ;  $F_2(2, 22) = 10.22$ ,  $MSE = 1,040,857$ ,  $p < .001$ . Comparisons between the different expression conditions, conducted with the Bonferroni–Dunn correction, revealed that the new expression condition was significantly longer than the old expression condition ( $ps < .0005$ ) and the first expression control condition ( $ps < .005$ ). The difference between the old and the first conditions was not significant ( $ps > .28$ ). The main effect of director knowledge was significant only in the item analysis,  $F_1 < 1$ ;  $F_2(1, 11) = 5.00$ ,  $p < .05$ ; there was no Director Knowledge  $\times$  Expression interaction,  $F_1(2, 70) = 2.32$ ;  $F_2(2, 22) = 1.80$ ,  $ps > .1$ .

If matchers interpret a new expression as implying a new referent, then they should look more at the temporary competitor object in the new condition. This is indeed what we found. The different expression conditions differed significantly both in terms of number of looks to the competitor object,  $F_1(2, 70) = 14.91$ ,  $MSE = 1.60$ ,  $p < .0001$ ;  $F_2(2, 22) = 6.00$ ,  $MSE = 0.96$ ,  $p < .01$ , and in terms of the duration of looking at the competitor,  $F_1(2, 70) = 19.68$ ,  $MSE = 594,424$ ,  $p < .0001$ ;  $F_2(2, 22) = 6.13$ ,  $MSE = 359,115$ ,  $p < .01$ . Comparisons conducted with the Bonferroni–Dunn correction revealed that matchers looked more at the target object in the new, compared with the old, expression condition (for both number and duration of fixations,  $ps < .005$ ). The difference between the new and the first expression conditions was significant in the participants analysis and marginal in the items analysis (number of fixations,  $p_1 < .01$ ,  $p_2 < .07$ ; duration of fixations,  $p_1 < .0001$ ,  $p_2 < .03$ ). The old and the first expression conditions differed only in the analysis by participants (for the number of fixations,  $p_1 < .01$ ; marginal for fixations duration,  $p_1 < .07$ ; all others, *ns*). There was no significant effect of director knowledge (for both number and duration of looks,  $F_s < 1$ ), nor was there a significant Director Knowledge  $\times$  Expression interaction: number of looks,  $F_1 < 1$ ;  $F_2(2, 22) = 1.81$ ,  $p > .18$ ; duration of looking,  $F_1(2, 70) = 1.53$ ,  $MSE = 46,350$ ,  $p > .2$ ;  $F_2 < 1$ .<sup>3</sup>

Consistent with the results of previous research (Metzing & Brennan, 2003) and of Experiment 1, matchers were slowed down when the speaker violated her own precedent. Because the comparison between the old and the new expression conditions involves a comparison across different referring expressions, and because the adjective–noun combinations used in the new expres-

sion condition were temporarily ambiguous before the noun region, it may be argued that the delay does not reflect the effect of the violation of a linguistic precedent but rather that the new referring expression is simply harder to understand, perhaps because it was temporarily ambiguous. However, a comparison of the new and the first conditions, which involved the same referring expressions, reveals this is not the case. Matchers took significantly longer to first fixate on the target object and to touch the target object in the new condition. This suggests that the delay in the interpretation of the referring expression was due to the violation of a precedent rather than to the temporary referential ambiguity of the expression. Thus, with the first expression control condition, Experiment 2 confirms the findings of Experiment 1. The clear cost associated with the violation of a precedent shows that listeners were indeed expecting the speaker to continue using the same expression even in the new context of the test grids. Because the first and the new expression conditions use the same expressions and the same referential contexts and differ only in terms of the existence of a linguistic precedent, the reliable difference between them suggests that the precedent indeed had an effect on listeners' expectations and on their interpretation of the new expression.

Matchers were not delayed in the first expression condition compared with the old expression condition, although we might have expected a larger benefit from the existence of precedents in the old expression condition (a benefit which would have been reflected in shorter response times in the old condition compared with both the new and the first expression conditions). Several factors may explain the lack of benefit for old expressions. First, as mentioned before, different expressions were used in the new and the old expression conditions. As a result, a difference may be due to an intrinsic difference between these expressions. For example, the old expressions may have represented better labels for the objects compared with the new expressions. Although we cannot rule out this explanation, it seems less likely given the fact that the

<sup>3</sup> An analysis excluding 7 additional participants (3 in the no-knowledge condition and 4 in the knowledge condition) who suspected the director was trained beforehand revealed the same pattern of results. There was a significant effect of expression on touch time,  $F(2, 54) = 21.41$ ,  $MSE = 3,724,607$ ,  $p < .0001$ ; first fixation to the target,  $F(2, 54) = 13.89$ ,  $MSE = 1,086,186$ ,  $p < .0001$ , and looks to the competitor, number of fixations,  $F(2, 54) = 12.38$ ,  $MSE = 1.32$ ,  $p < .0001$ ; duration of fixations,  $F(2, 54) = 18.14$ ,  $MSE = 615,226$ ,  $p < .0001$ . No other effects were significant ( $ps > .3$ ).

new expressions were temporarily ambiguous. The trend toward more looking at the competitor object in the first expression condition compared with the old expression condition (significant in the analysis by participants) suggests that matchers were sensitive to this ambiguity. Second, the lack of a difference between the first and the old expression conditions may not be entirely surprising given that matchers were highly familiar with the objects in the display at the time of the critical instruction and were able to pick out the intended referent relatively quickly even when they encountered new expressions. The lack of a difference in the latencies to the first and the last fixations on the target objects, despite the difference in the fixations on the competitor object, is consistent with this idea that matchers were able to quickly determine the intended referent. This finding is also consistent with the lack of a difference between original and new expressions produced by new speakers in Metzing and Brennan's (2003) study.

In contrast to the predictions of the cooperativeness-based account, matchers were delayed in interpreting referring expressions when the speaker violated her own precedents even when they believed that the knowledge regarding those precedents was not mutual. This pattern of results suggests that listeners expect the speakers to be consistent in their referring expressions even when such expectation is not motivated by an assumption of cooperativeness. Moreover, there was no evidence for a larger effect when consistency was implied by cooperativeness—that is, when listeners believed that knowledge regarding precedents was mutual—and accordingly, cooperativeness implies that the speaker should adhere to her own precedents.

This explanation relies on the assumption that the only relevant difference between the two director knowledge conditions involved matchers' beliefs about the mutuality of the precedents. As in Experiment 1, one potential concern is that although the critical instructions were the same across the two conditions, there may have been differences in the director's speech. For example, because the director in the knowledge condition actually produced the old expressions in the practice phase, they may have been less intelligible during the test phase (Bard et al., 2000; Fowler & Housum, 1987), consequently obscuring any extra benefit for the old expressions in the knowledge condition. To gauge such potential differences in the director's speech between the two conditions, we analyzed the director's utterances during the test phase. A coder, blind to the hypothesis of the experiment and to the conditions, measured the total duration of each critical instruction, the duration of the referring expression, and the number of hesitations during each instruction. Complete means are presented in Table 4.

We conducted a repeated-measures analysis of variance with knowledge as a between-subjects factor and expression as a within-subject factor. Analysis of hesitations during test instructions revealed no significant effects ( $F_s < 1$ ). Analysis of the duration of the total utterance showed only a significant effect of knowledge; overall, the director spoke faster in the knowledge condition compared with the no-knowledge condition,  $F(1, 35) = 5.52$ ,  $MSE = 3,620,462$ ,  $p = .03$ . Because the director in the knowledge condition was faster across all three expression conditions, it is safe to conclude that this increase resulted from practice in giving instructions rather than from repetition-induced reduction in word duration. This difference suggests that the overall greater

response times for touching the target object in the no-knowledge condition resulted from longer director's instructions.

Finally, analysis of the duration of the referring expression itself showed a Director Knowledge  $\times$  Expression interaction that, although not statistically significant, had a  $p$  level of  $< .08$ ,  $F(2, 70) = 2.70$ ,  $MSE = 11,422$ , and no significant main effects. Because shorter duration may be associated with a reduction in intelligibility, such an interaction may suggest that the numerically greater (though not statistically different) latencies of first fixation on the target for old expressions in the knowledge condition resulted from less clear expressions. If matchers in the knowledge condition are indeed delayed because of the director's less clear articulation, it may obscure the benefit for matcher from old expressions. Can the shorter duration of the old expressions in the knowledge condition compared with the no-knowledge condition account for the lack of a Director Knowledge  $\times$  Expression interaction? To evaluate this question, we analyzed a subsample of 30 participants selected to match in expression duration<sup>4</sup>, with the constraint that there would be an equal number of participants from each condition, balanced in terms of the test instructions version. If the difference in the duration of the referring expressions between the conditions is responsible for the pattern of results found for first fixations in the main analysis, we would expect the subsample of the data to not show the same pattern of results. Analysis of the subsample showed the exact same pattern of results found for the entire sample: There was a significant main effect of expression,  $F(2, 56) = 11.40$ ,  $MSE = 920,790$ ,  $p < .0001$ , and no significant effect of director knowledge ( $F < 1$ ) or Director Knowledge  $\times$  Expression interaction,  $F(2, 56) = 1.57$ , *ns*. Moreover, numerically, the data for the subsample showed the same trend toward greater latencies of first fixation in the knowledge condition compared with the no-knowledge condition (see Figure 2), suggesting that the main pattern of results is not due to differences in the duration of the referring expressions. This suggests that the greater latencies of first fixation in the knowledge condition do not reflect a penalty for shorter, and thus less clear, referring expressions.

## General Discussion

The experiments reported here investigated whether speaker specificity in comprehension is governed by cooperativeness or by a more general expectation of consistency. The results of both experiments suggest that the expectations are independent of cooperativeness. Listeners were delayed in identifying referents when the speaker was not consistent in her use of referring terms even when knowledge about the inconsistency was not mutual. Because this knowledge was not mutual, cooperativeness does not imply the speaker should be consistent. Therefore, cooperativeness cannot account for our results. Our results suggest that listeners' expectations are indeed speaker specific but not partner specific. They are directed toward a specific speaker, but they need not arise from assumptions of cooperativeness and accordingly are not restricted to a specific speaker–addressee dyad.

This kind of general expectation of consistency makes sense in the context of E. V. Clark's (1988, 1990) principle of contrast. If

<sup>4</sup> From analysis of expression duration, director knowledge,  $F < 1$ ; Director Knowledge  $\times$  Expression,  $F(2, 74) = 1.27$ , *ns*.

Table 4  
*Director's Utterances in Experiment 2*

Dependent variable	Knowledge						No knowledge					
	First		New		Old		First		New		Old	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Duration of expression (ms)	714	14	777	19	740	15	757	19	755	22	773	14
Total duration of utterance (ms)	4,938	153	5,087	150	4,962	132	5,238	115	5,526	194	5,306	173
No. of hesitations	1.12	0.19	1.08	0.17	1.13	0.20	1.26	0.15	1.33	0.14	1.10	0.14

the choice of a referring expression indicates a specific perspective the speaker adopts towards an object, listeners have a good reason to expect the speaker to hold the same perspective across conversational partners in the absence of a reason for a change. A change in referring expression can thus license certain inferences about the speaker's intended referent even if the change is across conversational partners. For example the change from "the silk shirt" to "that old thing" licenses certain inferences about the speaker even if the two expressions are used with different addressees. In the context of a referential communication task, referential success is the primary goal, and therefore it is reasonable that a change in form indicates a change in referent.

We do not claim that the expectation of consistency is never restricted to a particular conversational partner. This may be the case when an expression is clearly tailored for a specific addressee. For example if the speaker says "vest" instead of "undershirt" while speaking to a speaker of British English, listeners may not expect him to use the same expression while speaking to an American English speaker. In this context, listeners can infer that the main reason for using "undershirt" in the first place was the speaker's catering to the needs of the specific addressee. This may create an addressee specific precedent and could motivate the speaker's change to "vest" in subsequent interactions with Americans. People may also establish private lexicons with the purpose of not being understood by others. If you and I refer to a colleague as "Her Highness" and intend for others not to recognize to whom we refer, I may not expect you to use this expression when conversing with other people. Furthermore, it is possible that assumptions regarding speakers' cooperativeness or the specific communicative context can mediate the operation of listeners' expectation of consistency. Barring such special circumstances, our findings show that people expect speakers to keep their precedents independently of cooperativeness.

Our experiments do provide evidence for listeners' online use of their knowledge about the speaker in interpretation. Knowledge about the speaker's past expressions, whether or not it is mutual, is a source of contextual information that can constrain linguistic interpretation. Our results suggest a mechanism that is different in critical respects from Grice's (1975, 1989) idea of conversational implicatures. According to Grice, a speaker conversationally implicates that *p* only if the speaker expects the addressee to recognize that the speaker thinks that *p*. By using a different expression, a speaker implicates a change in intended referent only if the speaker expects the addressee to recognize the intention to signal such a change. Applying this to the situation in the present experiments, with such a change in expression, addressees can reason

that "if the speaker had wanted me to move the plastic cup again, assuming she is being cooperative, she would have said 'move the plastic cup.' Since she asked me to move the blue mug, she must have been talking about a different object." Reaching this inference would have required addressees to think that the speaker intended them to recognize that she had meant to implicate this. Thus, we can say that addressees inferred that the speaker implicated a change in referent by choosing a different expression only if they believed the speaker intended them to recognize the change in expression as implicating such a change. This is obviously not the case in the no-knowledge condition. Addressees could not have believed that speakers meant for them to recognize that they are using terms that are different from expressions they used before, given that as far as the speaker knows, the addressee has no idea what expressions they used in the past.

So the expectation of consistency could not have derived from an assumption of cooperativeness. Instead, we suggest that inferences made by listeners would be better characterized as situated inferences, in which listeners infer the intended referent based on the speaker's utterance and the situation in which it takes place, rather than as conversational implicatures that entail considering the speaker's communicative intentions (cf. Gauker, 2001). Listeners can make an inference and reach the same conclusion without invoking the assumption of cooperativeness and without viewing the speaker's consistency as addressee directed. Addressees do not even need to think that the speaker considered their previous usage in choosing the new expression: "The speaker called this 'the plastic cup' before; there appears to be no reason to change this expression when referring to the same object, the speaker is probably referring to a different object and is therefore using a different expression." Of course, this need not be an explicit conscious inference; listeners may not be aware of such an expectation of consistency. In this sense, interpretation of the speaker's use of referring expressions may be similar to the interpretation of behavior in general. In the absence of a discernible motivation, listeners would expect speakers to be consistent and would assume that a different expression picks out a different referent.

Speaker-specific effects on the comprehension of referring expressions suggest that listeners encode the identity of the speaker into the memory representation of the mapping between objects and referring expressions (Metzing & Brennan, 2003). Several models of memory (e.g., Gillund & Shiffrin, 1984; Hintzman, 1988; Humphreys, Bain, & Pike, 1989; Raaijmakers & Shiffrin, 1981) suggest that memory traces incorporate contextual information as well as item information and that the probability of remembering increases when the test context matches the context during

encoding. Thus speaker-specific effects may not reflect a conscious or unconscious goal-directed inference about speaker's linguistic behavior but may emerge as a result of normal memory processes in which the identity of the speaker acts as a retrieval cue for an expression-object mapping established by the same speaker during encoding. If the identity of the speaker is salient during encoding and acts as a relatively durable memory cue (compared with other contextual details), it can result in speaker-specific activation of existing expression-referent mappings. The activated mapping may then interfere with the establishment of a new mapping between the same referent and a different expression. A cue provided by the speaker's identity may be, however, swamped by a relatively much stronger cue provided by the repetition of the same expression, which may explain why a speaker-specific facilitation benefit was not found. In this way, effects that appear to be driven by special processes involving computation of common ground (such as speaker-specific effects) emerge in fact out of ordinary memory processes (see Horton & Gerrig, 2005). However, it is important to note that the speaker's identity serves as a memory cue for the established mapping independently of the assumption of cooperativeness and can act as a cue even for expression-object mappings that are not assumed to be mutually known.

Like others before, we have shown that listeners indeed expect speakers to adhere to their own precedents. The critical point here is that these expectations are governed not by an assumption of cooperativeness but by a general expectation of consistency. This expectation of consistency may reflect listeners' inferences about a speaker's current linguistic behavior on the basis of that speaker's past behavior; alternatively, such expectation may reflect ordinary non-goal-directed memory processes in which the speaker's identity acts as a retrieval cue for a stored expression-referent mapping. Knowledge of the speaker's past linguistic behavior affects listeners' interpretation of the speaker's utterances even when such knowledge is not mutual. I know you said it before, therefore I expect you to say it again.

## References

- Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, 42, 1–22.
- Barr, D. J. (2004a). The emergence of conventions in language communities. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 311–344). Cambridge, MA: MIT Press.
- Barr, D. J. (2004b). Establishing conventional communication systems: Is common knowledge necessary? *Cognitive Science*, 28, 937–962.
- Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory & Language*, 46, 391–418.
- Bloom, P. (2000). *How children learn the meaning of words*. Cambridge, MA: MIT Press.
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1482–1493.
- Brennan, S. E., & Metzing, C. (2004). Two steps forward, one step back: Partner-specific effects in a psychology of dialogue. *Behavioral and Brain Sciences*, 27, 192–193.
- Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, 19, 441–472.
- Clark, E. V. (1988). On the logic of contrast. *Journal of Child Language*, 15, 317–335.
- Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language*, 17, 417–431.
- Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge, England: Cambridge University Press.
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, 6, 813–839.
- Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, 37, 630–641.
- Eberhard, K. M., Spivey-Knowlton, M., Sedivy, J., & Tanenhaus, M. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, 24, 409–436.
- Fowler, C., & Housum, J. (1987). Talkers' signaling of "new" and "old" words produced in various communicative contexts. *Language and Speech*, 28, 47–56.
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27, 181–218.
- Gathercole, V. (1989). Contrast: A semantic constraint? *Journal of Child Language*, 16, 685–702.
- Gauker, C. (2001). Situated inference versus conversational implicature. *Noûs*, 35, 163–189.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 91, 1–67.
- Grice, P. H. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 225–242). London: Academic Press.
- Grice, P. H. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple trace memory model. *Psychological Review*, 95, 528–551.
- Horton, W. S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. *Cognition*, 96, 127–142.
- Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review*, 96, 208–233.
- Keysar, B. (1997). Unconfounding common ground. *Discourse Processes*, 24, 253–270.
- Keysar, B., Barr, D. J., & Lim, S. (2001, November). *Does conversational grounding involve partner-specific pacts? An example of the adaptive use of language*. Paper presented at the 42nd Annual Meeting of the Psychonomic Society, Orlando, FL.
- Markman, E. (1990). Constraints children place on word meaning. *Cognitive Science*, 14, 57–77.
- Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49, 201–213.
- Pickering, M. J., & Garrod, S. (2004). Towards a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–225.
- Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, 88, 93–134.
- Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21, 211–232.
- Sedivy, J. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research*, 32, 3–23.



Appendix

Original and New Expressions for Target Items

Original expression	New expression
Sunglasses	Broken glasses
Elephant rattle	Baby rattle
Fire engine	Red truck
Sheriff's badge	Gold star
Orange block	Plastic cube
Polar bear	Keychain bear
Yellow car	Toy jeep
Ceramic cup	Blue mug
Lego bridge	Multicolored gate
Plastic cup	Measuring cup
Shiny bow*	Silver bow
Stuffed snake*	Pink snake

*Note.* Expressions indicated with an asterisk were used only in Experiment 2.

Received May 30, 2006

Revision received November 6, 2006

Accepted November 6, 2006 ■

## Low Publication Prices for APA Members and Affiliates

**Keeping you up-to-date.** All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*. High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

**Essential resources.** APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association*, and on dozens of new topical books each year.

**Other benefits of membership.** Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

**More information.** Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.