

---

# Spatial Dialogue between Partners with Mismatched Abilities

MICHAEL F. SCHOBER

## 3.1 Introduction

A fundamental fact of spatial dialogue is that speakers can have different perspectives on what they are talking about: what is on one speaker's right can be on the other's left, or in front or behind (in back in US English), for the other. Of course, this is part of a larger truth about dialogue: beyond having different views of what is in front of them (their heads can never occupy precisely the same space, unlike Steels and Loetzschi's robots, this volume), speakers can have different world views, different agendas, different conceptualizations of the topics at hand, and different beliefs about what their conversational partner believes (see, for example, Russell and Schober, 1999; Schober, 1998a, 2005). Speakers thus confront the philosopher's notorious 'problem of other minds' every time they talk with each other, whether they know it or not, at multiple levels. At the level of spatial language, every use of a locative expression reflects a choice of reference frame or coordinate system, and thus provides evidence about how the speaker has attempted to solve the other minds problem at the current moment.

To consider this in a concrete setting, imagine that two people are facing each other over dinner at a restaurant. Jeff arrived early and ordered two different kinds of white wine, reflecting their different tastes. Aimee arrived later to find two similar-looking filled glasses on the table. The glass that is on Jeff's right is on Aimee's left. If Aimee asks which glass has the Riesling in it, Jeff has a number of choices, each of which reflects a different solution to the other minds problem. Jeff could avoid the problem by pointing silently at one glass, or pointing and saying 'that one.' But if he would rather not point in public, he must rely on language. He can say it's 'the one on your right', which not only uses Aimee's reference frame but marks it as Aimee's with the term 'your'. He can say that it's 'the one on the right', which takes Aimee's perspective but—riskily—assumes much more about Aimee's mental state: he assumes that Aimee will know that it must be Aimee's right that is being referred to. He can say it's 'the one on the left,' which takes a different risk in assuming that Aimee will know that Jeff is using his own perspective, or 'the one on my left,' which explicitly marks whose perspective is being used.

Of course, Jeff has the insurance plan that comes with participating in a dialogue (and not monologue); if Riesling-seeking Aimee has any doubts about whether she has understood the description she can ask for clarification ('Which one? or 'You mean *your* left?').

Or consider a slight variation: the dinner companions are seated at a corner booth, with a 90-degree discrepancy in their viewpoints on the scene. Depending on how the glasses are laid out on the table, the terms 'right' and 'left' might well work from both parties' perspectives (what I have called a 'both-centered' perspective, Schober, 1993), and so there might be less need to mark whose left or right is being described. Other arrangements of objects and dinner companions will afford different potential descriptions, and will bring along with them ways of marking how much is assumed about what the partner knows. If, for example, there were three glasses as potential referents, Jeff could say that it is the one 'in the middle,' which would be true no matter where he and Aimee were sitting; this sort of 'neutral' perspective (Schober, 1995) can allow speakers to avoid having to choose one or the other's perspective. If any objects on the dinner table have their own coordinate system (front, back, top, bottom, left, right)—say, an unusual breadbasket, or a model plane—Jeff could use an object-centred perspective to refer to the desired glass as behind or to the left of that object.

Now, we already know that a number of factors affect spatial perspective choice. As I just noted, different physical settings afford different perspectives—objects' relative locations and the nature of the objects themselves (whether or not they have their own coordinate systems) allow different kinds of spatial descriptions (e.g., Levelt, 1982; Ullmer-Ehrlich, 1982; Wunderlich, 1981). The degree of offset of the partner's vantage point can affect a speaker's perspective choice (e.g., Bürkle, Nirmaier, and Herrmann, 1986; Herrmann, 1989; Bürkle, and Nirmaier, 1987). The speaker's social goals (politeness, respect, egalitarianism) can affect perspective choice: college students are more likely to take a professor's spatial perspective than a fellow college student's (e.g., Graf, described in Herrmann, 1989), and people are likely to take their partner's perspective just as much as their partner has taken theirs (Schober, 1993). In earlier work (Schober, 1993, 1995, 1998b), I have shown that conversational feedback can affect perspective choice—people are more likely to take the perspective of a (silent) imaginary partner than of a live partner who can give feedback of understanding. And people are likely to stick with perspectives already taken with their partner.

Here I argue that there are additional differences between dialogue partners that can affect spatial perspective choice. The more I have thought about the issues, the more negotiating spatial perspective in dialogue strikes me as hard to divorce from negotiating the rest of what speakers negotiate in dialogue: their agendas (small talk or serious talk?) and thus their need for precision, their likelihood of requesting clarification, their desired level of politeness, and their assessments of each other's relevant abilities. Speakers have been observed to accommodate to each other on multiple dimensions, and there are many unanswered questions

about how this accommodation works and how individually variable it is (see Schober and Brennan, 2003, for discussion). Here I will focus on one of these factors—mismatched spatial abilities—and describe some results from a study (Schober, 1998c) that demonstrates how ability mismatches can drive perspective choice and affect the dialogue more generally.

### 3.2 Study

Why should spatial abilities affect linguistic perspective choice? At a fundamental level, taking one's partner's spatial perspective linguistically requires being *able* to see (or at least describe, which may or may not be the same) things from that point of view. How easy or hard this is to do turns out to vary across individuals (and along multiple dimensions which are correlated—see Hegarty and Waller, 2004, among others). If Aimee finds it extremely easy to imagine what Jeff is seeing, then it should be particularly easy for her to use language that reflects Jeff's coordinate system when describing a location for Jeff, and she should find it particularly easy to understand a description from Jeff's point of view. If Jeff has a terrible time imagining what Aimee is seeing, then it should be particularly difficult for him to produce and to comprehend a spatial description from Aimee's point of view. When Aimee and Jeff meet up, their relative abilities are, I propose, likely to affect what happens in the dialogue.

This hypothesis is lent plausibility by earlier evidence (Graf, cited in Herrmann, 1989) that college students are more likely to take the perspective of an (imaginary) small child than that of a fellow college student, presumably because they are less confident that the child will be able to understand descriptions that don't take their point of view. It is consistent with the finding that when speakers lack evidence about whether their partners have understood their spatial descriptions, they are more likely to take the addressee's perspective (Mainwaring et al., 2003; Schober, 1993). The hypothesis is also lent plausibility by the many demonstrations of how speakers adapt to their less capable partners in other arenas. Normally abled caretakers adapt to their conversational partners with mental retardation on several dimensions (Abbeduto, Weissman, and Short-Meyerson, 1999; Testa, 2005). Adults with greater expertise tailor their utterances for their partners with lesser knowledge (Isaacs and Clark, 1987), and even children adapt their speech to younger children with less linguistic ability (Shatz and Gelman, 1973).

The experiment described here tests this hypothesis by asking people to describe locations for partners. Unbeknownst to them, participants had been selected for having very high or very low mental rotation abilities as assessed by performance on a timed mental rotation test, the Card Rotations Test (S-1 rev.) from the ETS Kit of Factor-Referenced Cognitive Tests (19/6/1992). Of course, the abilities that a mental rotation test taps are disputed; it is not clear whether test-takers mentally rotate the figures or themselves—see, for example, Just and

Carpenter, 1985; Wraga, Creem, and Profitt, 2000; Zacks, Mires, Tversky, and Hazeltine 2000—or whether they carry out imagistic or propositional or some other sort of transformation—see, for example, Polyshyn, 2002. But the intuitive notion here was that whatever abilities a mental rotation test taps are likely to be necessary for spatial dialogue when partners have different points of view on a scene.

**Participant selection.** Participants were not informed about how they had been selected for the study; the preselection test was part of a packet of questionnaires given long before the study to over 700 students in large Introductory Psychology classes at SUNY Stony Brook. The test consisted of two sets of 80 items which ask whether two simple figures at different rotations are the same or different (reversed); students were given three minutes to attempt each set of 80. Scores on the test are typically normally distributed. For the purposes of this study, students who averaged above 54 out of 80 on the two tests were considered to be of high ability, and students who scored below 37 out of 80 were considered to be of low ability.

**Experimental task.** Based on these scores, 70 pairs of students were brought to the laboratory. In each pair, one student, randomly selected to be the *director*, was seated in front of a computer monitor; his or her task was to describe locations on a series of 32 displays (presented in different random orders for different pairs), which were designed so that each locative description could unambiguously be coded as reflecting only one perspective (round 1). The other student, the *matcher*, marked locations in a paper packet with the same 32 displays. After doing this, the students switched roles and described 32 more displays (round 2).

The displays were designed to require participants to take one or another perspective from among five possible perspectives as described in Schober (1995): speaker-centred, addressee-centred, both-centred, object-centred, and neutral. Speaker-centred descriptions are true from the speaker's point of view and not from the addressee's, and addressee-centred descriptions are true from the addressee's point of view and not the speaker's (whether or not they are explicitly marked for speaker, as in 'my left' or 'your left'). Both-centred descriptions are true from both the speaker's and the addressee's point of view (e.g., 'on the left' when both parties are looking from more or less the same angle) and are not otherwise marked for which speaker's perspective is intended; only some arrangements of speakers allow both-centred descriptions, but facing one another does not. Object-centred descriptions (or 'intrinsic' descriptions; see Watson *et al.*, this volume) reflect the coordinate system or metaphorical point of view of a non-human object that has its own front and back or left and right, as automobiles and other vehicles do; of course, such descriptions may reflect the imagined point of view of a human using the vehicle, but the critical point is that they reflect the object's orientation and perspective independent of the human observers (that is, no matter where speakers are in relation to the car or each other, the exhaust pipe is still at the back end of the car). And neutral descriptions are other descriptions

that are independent of the human observers' points of view but that do not reflect the intrinsic perspectives of the objects in the scene; for example, if a target object is 'between' two cars, this will be the case no matter which direction the cars or the observing humans are facing.

As in Schober (1993), the displays were minimal and simple, consisting of objects without their own coordinate system (circles) in various arrangements, sometimes including an additional simple object with its own clear front, back, left, and right (an aeroplane). The idea was to create a set of stimuli (1) with the minimum complexity necessary for distinguishing between these different perspectives; (2) that contrasted addressees and objects in four different rotations (0, 90, 180, 270 degrees of disparity); (3) that alternated which circle was the target object, and (4) that allowed for unambiguous descriptions that could clearly be coded for one perspective only. Obviously, these scenes substantially underestimate the complexity and ambiguity of possible scene arrangements in the real world, but they represent a reasonable set for experimental purposes.

The final set of 32 displays (sampled from the 61 logically possible displays from our combination of elements) included eight two-circle displays that required partners to use either the speaker's or addressee's perspective (see Figure 3.1a); eight that allowed speaker-centred, addressee-centred, and neutral description (see Figure 3.1b); eight that allowed speaker-centred, addressee-centred, and object-centred descriptions (with the addition of the outline of an aeroplane) (see Figure 3.1c); and eight that allowed not only speaker- and addressee-centred but also neutral and object-centred descriptions (see Figure 3.1d).

To exemplify, in Figure 3.1a speakers must refer to the target circle either as the one on the right or the left, which reflects a choice of one party's perspective ('right' reflects the speaker's perspective and 'left' reflects the addressee's); displays in this series were arranged so that there were never indeterminate views that might allow both-centred descriptions. In Figure 3.1b, the speaker could refer to the target circle as the one in front of the others (speaker's perspective), behind (in back of) the others (addressee's perspective), or the one in the middle (neutral perspective). In Figure 3.1c, the target circle is on the left (speaker-centred), on the right (addressee-centred), or behind the plane (object-centred description). And in Figure 3.1d the target circle is in the left row of circles, the second one back, from the speaker's perspective; in the back or far row, second from the left or right, from the addressee's; on the right from the plane's (object-centred) perspective; and in the middle circle in the row of three or next to the wing (true from any perspective, and thus neutral). To locate the target circle in this more complex scene, the speaker could easily use more than one locative description, potentially reflecting more than one perspective, as in 'it's the middle circle (neutral) on my left (speaker-centred)'. In this scene a speaker could even use a both-centred perspective if they were to refer to the bottom-left circle as the circle 'on the left', which is true from both the speaker's and addressee's point of view, and indeterminate as

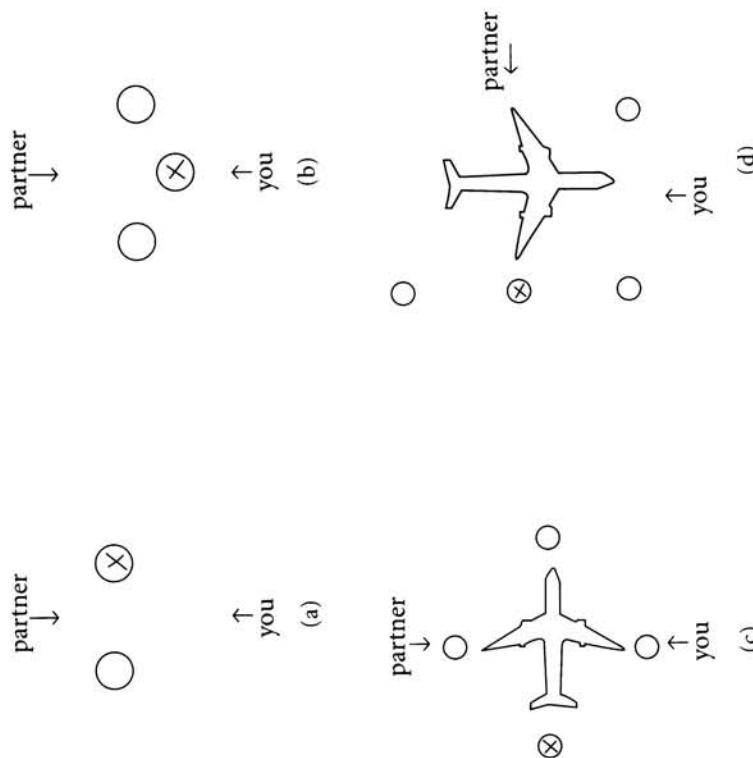


FIG. 3.1. Sample stimuli

to which it is. But the target circles were selected so as to avoid such indeterminacies. So in some of the displays there were as few as two possible perspectives for participants to use (director's vs. matcher's), and in some there were as many as five.

After the experiment, each participant again took the timed mental rotation test to verify their ability level, because low scores in preselection could reflect laziness or inattention rather than actual ability. After eliminating the surprisingly high number of pairs whose scores were inconsistent with the earlier test, the conversations of the 29 remaining pairs were transcribed and coded for perspective. In this final sample, the mean score of high ability participants was 72.1, ranging from 54.5 to 80, and the mean score of the low ability participants was 19.2, ranging from -3.5 to 36.5 (a negative score meant that the participant went slowly and got more comparisons wrong than right).

*Transcription and coding.* Audiotapes of the conversations were transcribed. Each locative description was coded by two coders (the few disagreements were resolved by discussion) for whether it reflected the director's perspective, the

matcher's perspective, a both-centred perspective (true from both interlocutors' points of view and not marked as belonging to one or the other), an object-centred (the plane's) perspective, or a neutral perspective (true from anyone's point of view). Any descriptions that did not reflect any interlocutor's (or any discernible) point of view were coded as bizarre.

Note that the taxonomy of perspectives reflected in this coding scheme differs from the more frequently employed distinction between intrinsic and relative perspectives, which in most cases treats all observer-centred perspectives as equal rather than distinguishing between speaker- and addressee-centred perspectives (although see Tenbrink, 2007, for schemes that make finer-grained distinctions). The taxonomy is also task-specific, in that the stimuli in this experiment—both the target objects to be described and the objects as arranged—were selected so that there would never be a possible descriptive scheme for target objects that was ambiguous with respect to these perspectives. In real-world scenes it is regularly possible for speakers and objects to be situated in ways that allow descriptions that are ambiguous with respect to perspective; if, for example, the plane in Figure 3.1c were facing one of the partners and the target circle were in front of or behind the plane, the description 'behind the plane' would be ambiguous with respect to whether it reflected the plane's object-centred (intrinsic) back or one of the interlocutors' points of view. With the arrangement in Figure 3.1c, the description 'behind the plane' unambiguously reflects an object-centred frame of reference (which holds for both speakers), while a description of 'on the (my) left' or 'on the (your) right' would reflect one of the interlocutors' points of view. Similarly, for Figure 3.1d, a description of the target circle as 'to the left of the plane' would reflect the speaker's perspective (as seen from 'you'), while 'behind the plane' would have to reflect the addressee's perspective, and not an object-centred (plane's coordinates) perspective, because the target circle is not at the plane's own back end. So it is the particular arrangements of interlocutors, display objects, and circles selected as targets that allowed for the unambiguous coding. (Of course, this also affects the generalizability of the results, in that it is unknown how frequently real-world settings allow interlocutors, or researchers, to set perspectives so unambiguously).

Note also the importance of the displays being presented within an experimental task that made comprehension unambiguously measurable, and independent of any coding or judgement of the suitability of locative expressions (see, for example, Carlson and Hill, this volume; Tenbrink, 2007, and this volume, for discussion of the communicative principles that guide object reference in contrastive scenarios). The criterion for comprehension accuracy was whether the matcher marked the appropriate circle on the display, whether or not the director's description might be considered ambiguous by an outside observer; in many other domains interlocutors come up with schemes for understanding each other that are not transparent to outsiders (see, for example, Schober and Clark, 1989), and I would argue that there is an important distinction to be made between codability

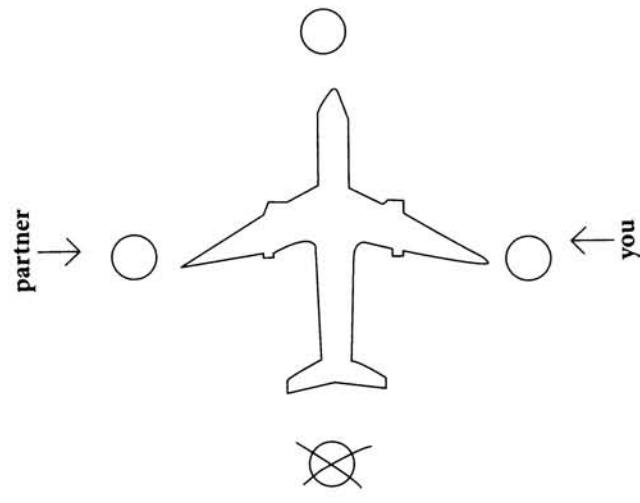


Fig. 3.2. Matcher's point of view on display, Round 2 #1

of a description as an overhearing outsider and comprehension accuracy for an addressee.

*Results.* What was immediately apparent in the transcripts was how much the partners' relative mental rotation abilities affected the nature of the interaction, despite the fact that participants were entirely unaware that they had been selected on the basis of their relative abilities. Compare the following exchanges, which were the first interchanges in the second round of interaction, after the interlocutors already had experience with the other in the alternate task role; in both cases the matcher saw the display in Figure 3.2.

The first is a typically brief interchange between a high-ability director paired with a high-ability matcher:<sup>1</sup>

D: It's behind the plane (OBJECT-CENTRED)

M: Okay

The matcher selected the correct target circle, despite the fact that she could have interpreted this as one of the other circles; clearly this pair had come to agreement about the use of object-centred perspectives during their earlier interactions.

The second exchange, describing the same display, is a far more tortuous interchange between a high-ability director and a low-ability matcher:

D: Okay my plane is pointing towards the left (DIRECTOR-CENTRED)

M: Okay

D: And the dot is directly at the tail of it (NEUTRAL)

D: Like right at the back of it (OBJECT-CENTERED)

M: Okay mine is pointing to the right (MATCHER-CENTRED)

D: Oh yours is pointing to the right (MATCHER-CENTRED)

M: Yeah

D: So your dot should be on the left (MATCHER-CENTRED)

D: Because my dot is on the right (DIRECTOR-CENTRED)

D: In back of it (OBJECT-CENTRED)

D: So your dot should be at the left (MATCHER-CENTRED)

D: At the back of it right (OBJECT-CENTRED)

M: Yeah

D: Yeah

M: But if it is the same—but if it—the same dot-right? Wait a minute, if my—  
your plane is pointing to the left \* [something]—\*

D: \*My\* plane is pointing to the left (DIRECTOR-CENTRED)

M: Mm-hm

M: And that dot and the dot that's highlighted is the one all the way in the back  
of it (OBJECT-CENTRED)

M: Like behind the tail (OBJECT-CENTRED)

M: Yes, so so my dot is gonna be

D: So my dot is on the right (DIRECTOR-CENTRED)

D: And yours should be on the left right (MATCHER-CENTRED)

M: Yeah

D: Okay \*so your—\*

M: \*Right behind the tail\* okay (OBJECT-CENTRED)

D: Okay

Simple counts of the numbers of words that each pair needed to accomplish their task demonstrate that directors in high-high pairs used marginally fewer words than directors in either mixed-ability or low-low pairs,  $F(1,26) = 3.60$ ,  $p < .07$  (see Figure 3.3).

Another striking finding was that low-ability interlocutors were far more likely to produce bizarre uncodable descriptions like this one, produced by a low-low ability pair for the display in Figure 3.4:

D: Um the top circle (BIZARRE)

M: There is no top

D: Oh wait um the bottom circle (DIRECTOR-CENTRED)

M: There's no bottom

<sup>1</sup> In the transcript excerpts, D refers to the Director and M to the Matcher. Overlapping speech is enclosed in asterisks and pauses are indicated by periods between spaces. Coded perspectives are indicated in capitals.

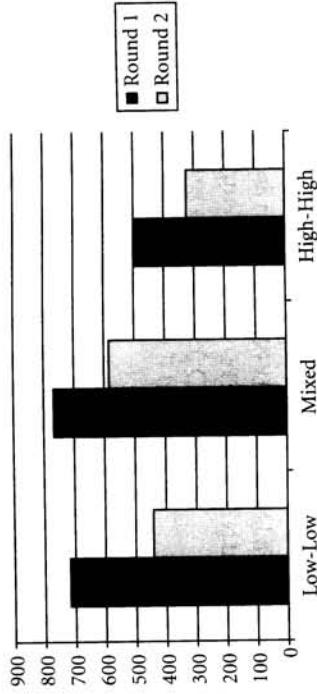


FIG. 3.3. Mean number of words spoken by directors by group

D: Oh the circle to the left left (BIZARRE)

M: Okay

Note that the director seems to be stabbing in the dark, hoping for a lucky description that the matcher will find acceptable; note also that the low-ability matcher accepts a description that simply isn't true from either party's perspective and ends up marking the wrong circle. As Figure 3.5 shows, low-low pairs produced more bizarre descriptions than mixed pairs, and mixed pairs produced more than high-high pairs, linear trend  $F(1,26) = 5.74$ ,  $P < .03$ .

This stabbing in the dark pattern, and the seeming inability to decentre from their own points of view, was evident throughout the low-ability pairs' conversations, as in this example (see Figure 3.6):

D: The top circle (BOTH-CENTRED) to the right (DIRECTOR-CENTRED)

M: There's no top circle at the right

D: Uh, oh the bottom the bottom circle (BIZARRE) to the left of of (MATCHER-CENTRED) the picture

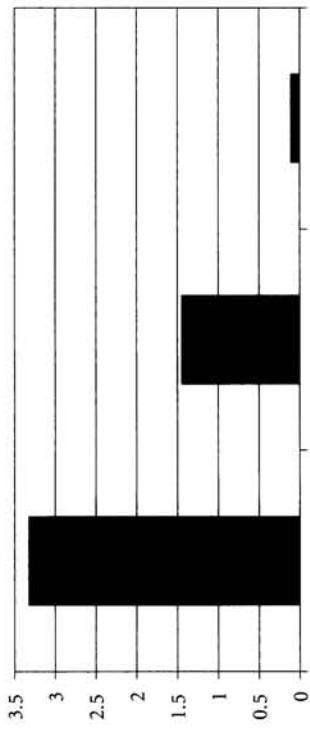
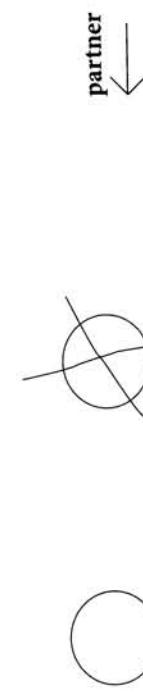


FIG. 3.5. Mean number of bizarre descriptions by group, Round 1

M: Um there's none at the left  
 D: Oh um the bottom circle (BIZARRE) to the right (DIRECTOR-CENTRED)  
 M: There's none at the right  
 D: Uh top circle (BOTH-CENTRED) to the left (MATCHER-CENTRED)  
 M: Okay

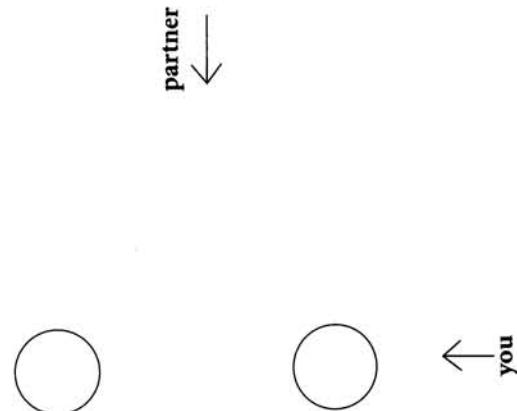


FIG. 3.4. Matcher's point of view on display, Round 1 #14

FIG. 3.6. Matcher's point of view on display, Round 1 #29

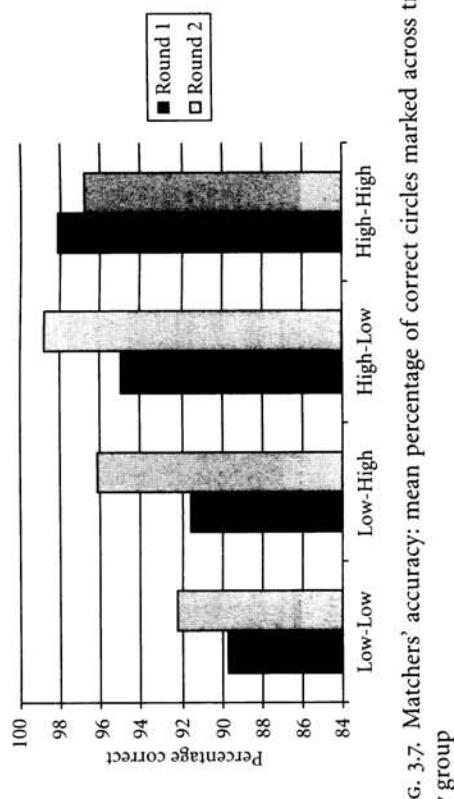


FIG. 3.7. Matchers' accuracy: mean percentage of correct circles marked across trials by group

Not surprisingly, matchers in low-ability pairs were rather inaccurate in marking the appropriate objects, indicating their generally poorer comprehension (see Figure 3.7). But the only reliable difference was between low-low and low-high pairs,  $F(1,25) = 4.47$ ,  $p < .05$ ; the comprehension for matchers in mixed pairs seemed to be protected by having one high-ability partner—whether matcher or director.

What about taking the other person's perspective? The very clear finding is that low-ability directors were more likely to take their own perspective, while high-ability directors were more likely to take the matcher's perspective, interaction  $F(1,25) = 9.31$ ,  $p < .005$  (see Figure 3.8). If we examine matcher-centred descriptions in further detail, we see that their use changed over the course of the 32 descriptions. As Figure 3.9 shows, if we compare the first eight and the final eight

descriptions, we see notable changes in the use of matcher-centred descriptions, interaction  $F(1,25) = 11.21$ ,  $p < .005$ , as if partners were getting to know each other's abilities as the round went along. Already in the first quartile, high-ability directors were more likely to take their partner's perspective. But as time wore on in mixed pairs, high-ability directors with low-ability matchers drastically increased their use of matcher-centred perspectives. Low ability directors with high-ability partners notably decreased their attempts at matcher-centred perspectives.

Closer examination of the transcripts suggests part of what is going on: high-ability directors recognize that their partners are having trouble understanding director-centred descriptions and switch to taking their partner's point of view more and more. And low-ability directors end up being licensed by their high-ability matchers' questions to use more director-centred descriptions, as in this example describing the display in Figure 3.10:

- D: Okay now, it's uh the second one (BOTH-CENTRED)
- M: The one towards the bottom of the page? (MATCHER-CENTRED)
- D: I'm sorry?
- M: The one towards the top of the page (MATCHER-CENTRED)
- M: Or the top? (MATCHER-CENTRED)
- D: Uh there are two circles near
- M: I oh yours are right next to each other (DIRECTOR-CENTRED)
- D: Yeah
- M: All right the one towards
- D: It's
- M: The left (DIRECTOR-CENTRED)
- M: Or right? (DIRECTOR-CENTRED)
- D: Left (BIZARRE)

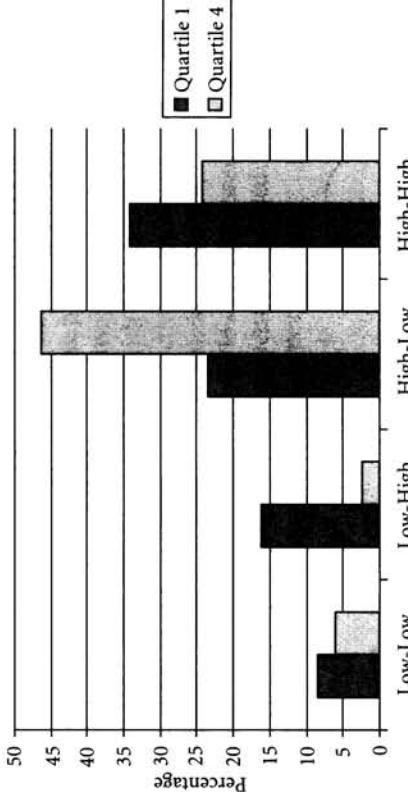


FIG. 3.9. Matcher-centred descriptions, Round 1

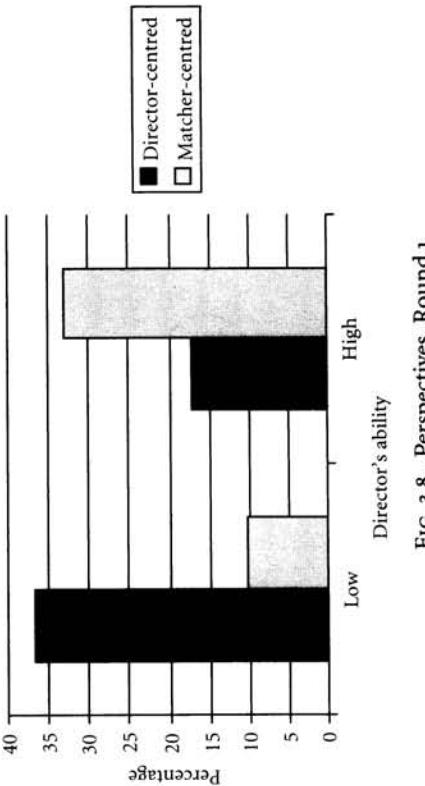


FIG. 3.8. Perspectives, Round 1

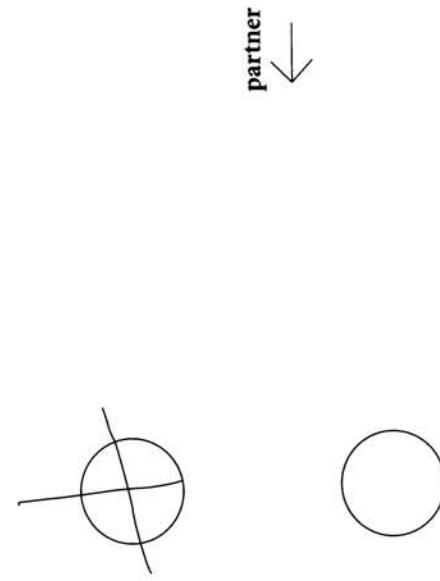


FIG. 3.10. Matcher's point of view on display, Round 1 #21

M: All right \*got it\*  
D: \*Okay\* mm-hm  
M: Yeah

Directors' and matchers' abilities did not affect their use of neutral and object-centred descriptions; the only findings with these perspectives were that directors of all abilities were more likely to use neutral and more object-centred descriptions at 90- and 270-degree offsets than when they were face to face, replicating the general pattern in Schober (1995) with a different set of displays.

Overall, what emerges from these findings is not only the extent to which the shape of the conversation is affected by individual abilities but also the extent to which the conversational context—the partner's uptake and input—shapes what each individual does. As has often been argued elsewhere, language in interaction is simultaneously an individual-mind and collective phenomenon, with social processes affecting individual cognition and individual cognition affecting social processes. Either level of analysis without the other leaves out something important.

More particularly, the findings demonstrate that not only individuals' spatial abilities but also their interlocutors' abilities can substantially affect people's choice of perspectives in describing locations, and that they subtly judge

each other's abilities within a few moments of conversing, by as yet unknown mechanisms. Pairs in which both partners had poor mental rotation abilities understood each other's spatial language more poorly than pairs in which at least one partner had high mental rotation ability, which seemed to allow pairs to compensate for the difficulties that people with low ability had. People with low ability were far more likely to provide ineffective or uncodable descriptions that were not true from anyone's perspective. Speakers with high ability were more likely to take the perspective of partners with low abilities, essentially encouraging them to speak egocentrically, and this propensity increased over time, as they gained further evidence of their partner's ineffectiveness.

### 3.3 Further Questions

Adding individual spatial abilities into theories of spatial language and dialogue raises a host of unanswered questions, just as the addition of individually variable cognitive capacities (like working memory span) has complicated theories of language processing more generally (see e.g. Kurtz and Schober, 2001; Miyake, 2000, among many others). First, these preliminary findings suggest that perspective choice in spatial language use is closely related to other aspects of spatial cognition—but which aspects and when? Given that different abilities, like spatial orientation and spatial visualization, are at least partially dissociable (Hegarty and Waller, 2004), how might the different abilities affect spatial language choice—and which aspects? At least under certain circumstances—certain spatial scenes, certain relative locations of partners—the use of spatial language by definition requires assessing the partner's point of view, and this requires some form of mental rotation. Might different rotation strategies that speakers take affect their spatial language use? The current findings only begin to suggest that the primary site of any effects will be in person-centred (speaker and addressee) perspectives, but this would obviously need to be studied with a greater range of scenes and in more naturalistic settings.

Second, much is unknown about how partners assess each other's abilities. The data here hint that a partner's failure to produce and comprehend sensible descriptions tips off high-ability speakers, but how does this work? Are particular kinds of failures particularly informative? Beyond varying in their abilities to produce and comprehend spatial descriptions, do speakers also vary in sensitivity to social cues, or empathy, or interest in perspective taking, in ways that interact with their spatial language choices? Ability assessments in other domains appear to be relatively quick, just as they were here: consider Isaacs and Clarks' (1987) New Yorkers who very quickly changed their level of detail about New York landmarks when they figured out their partners were non-New Yorkers. And there is some evidence that people can use the *way* their partners talk—the extent of pausing and disfluency, and doubtful tone of voice—to judge how confident their partners

are about what they are saying or how likely their partners are to need clarification (e.g. Brennan and Williams, 1995; Schober and Bloom, 2004). But is the process of assessing each other's relative abilities the same in spatial language as in other settings? And do people accommodate in the same ways?

Third, the current study selected participants whose abilities were on the extreme ends, and the differences in ability in mixed pairs were extreme. How do abilities come into play for less extreme abilities or for less extreme ability differentials? Note also that, because this study's participants' abilities were extreme, they would be likely to perform just as well or as poorly at the various correlated components of perspective taking. But to the extent that components are dissociable, how might those different components affect spatial language use?

Fourth, the findings here are necessarily limited in generalizability by the choice of stimuli and experimental setting. The stimulus displays were selected to allow unambiguous coding of perspective and clear measurement of the accuracy of description, and so the arrangements of scenes, stationary interlocutors, and target objects were carefully chosen, with recognition that they were odd. The experimental setting had the additional oddness of both parties' not seeing each other's faces, and looking at different representations of the same scene (one on a monitor, one on a piece of paper). Exactly how this all affected results is unknowable, and I would argue for future research that sets up situations that are closer to real-world settings. But this leads to the question of what real-world settings are really like and how they vary for different people and different realms of experience: as far as I am aware, it is entirely unknown how often speakers describe location in settings with so clearly definable perspective choices, how often the interlocutors remain stationary and can see each other's point of view, how often they use language (as opposed to pointing) to refer to target objects, and how often their descriptions allow their partners unambiguous access to whose perspective is being chosen.

Despite these limitations of the current study, I propose that spatial language provides an opportunity for researchers in other sorts of language and dialogue, because (in the right circumstances) the terms that interlocutors use allow clear assessment of who has taken whose perspective, what references speakers have intended, and what addressees have understood; in addition, researchers can manipulate factors that affect perspective-taking. This is an unusual opportunity for researchers interested in the problem of other minds who have focused on other sorts of dialogue; in other settings quite often one cannot tell whether people using the same words really have the same representations underlying them (Schober, 2005, 2006). That is, when two people are discussing 'abortion' or 'euthanasia' or even less loaded topics like their 'jobs', the mere fact that people use the same words is not a guarantee that they have the same underlying representation; lexical alignment can actually mask undetected conceptual misalignment, which can lead to miscomprehension and task failure (see, for example, Conrad and Schober, 2000; Schober, Conrad, and Fricker, 2004, for evidence from survey

interviews). The facts of spatial perspective actually provide a clear case where sometimes one person's use of 'left' means something quite different from the other's. This allows for empirical testing of claims about the extent to which two parties' conceptual alignment is automatic when the same words are used (Pickering and Garrod, 2004; Vorwerg, this volume; Watson *et al.*, this volume). As I have argued elsewhere (Schober, 1995, 1998a), it also allows for empirical testing of claims about the extent to which partners allocate effort within a pair as opposed to the individual, as in Clark and Wilkes-Gibbs' (1986; see also Clark, 1996) proposals about least collaborative effort, in a situation where each party's individual effort is definable.

### Acknowledgements

This work was supported by NSF Grants SBR-97-30140, IRI-94-02167, and SES-05-51294, as well as a New School Faculty Development Award. I am grateful to Susan Brennan and the SUNY Stony Brook Psychology Department; to Jon Bloom, Silvia Leon, Marcy Russo, Catherine DiNardo, Danielle Barry, Yassi Gidfar, Jack Shelle-Tremblay, and Susan Berrill for their help in recruiting and running participants and coding the data; and to the anonymous reviewers and the editors for their useful comments.